

A Counting Lemma for Binary Matroids and Applications to Extremal Problems

Sammy Luo

November 1, 2016

Abstract

In graph theory, the Szemerédi regularity lemma gives a decomposition of the indicator function for any graph G into a structured component, a uniform part, and a small error. This result, in conjunction with a counting lemma that guarantees many copies of a subgraph H provided a copy of H appears in the structured component, is used in many applications to extremal problems. An analogous decomposition theorem exists for functions over \mathbb{F}_p^n . Specializing to $p = 2$, we obtain a statement about the indicator functions of simple binary matroids. In this paper we extend previous results to prove a corresponding counting lemma for binary matroids. We then apply this counting lemma to give simple proofs of some known extremal results, analogous to the proofs of their graph-theoretic counterparts, and discuss how to use similar methods to attack a problem concerning the critical numbers of dense binary matroids avoiding a fixed submatroid.

1 Introduction

In this paper, the term *matroid* refers to a simple binary matroid. A *simple binary matroid* M is, for our purposes, a full-rank subset of $\mathbb{F}_2^r \setminus \{0\}$ for some positive integer $r = r(M)$ called the *rank* of M . The *critical number* $\chi(M)$, another important quantity, is the smallest c such that there is a copy of \mathbb{F}_2^{r-c} in $\mathbb{F}_2^r \setminus M$, or equivalently such that M is contained in a union $A_1 \cup \dots \cup A_c$ where each A_i is a hyperplane $\mathbb{F}_2^r \setminus \mathbb{F}_2^{r-1}$.

Basic examples of matroids include the following.

- The *projective geometry* of rank r , $PG(r-1, 2) := \mathbb{F}_2^r \setminus \{0\}$, which has rank r and critical number r .
- The *affine geometry* of rank r , $AG(r-1, 2) := \mathbb{F}_2^r \setminus \mathbb{F}_2^{r-1}$, which has rank r and critical number 1.
- The *Bose-Burton geometry*, $BB(r, c) = \mathbb{F}_2^r \setminus \mathbb{F}_2^{r-c}$, a generalization of both examples above, which has rank r and critical number c .

There is a direct connection between graphs and matroids: For any graph G we can define its *cycle matroid* $M(G)$, whose elements correspond to edges of G , where a set of elements of M is linearly independent if and only if the corresponding edges of G contain no cycle. A matroid is called *graphic* if it is the cycle matroid of some graph. It is easy to see that $\chi(M(G)) = \lfloor \log_2(\chi(G)) \rfloor$, where $\chi(G)$ is the chromatic number of G .

The critical number of a matroid is analogous to the chromatic number of a graph. Just as the chromatic number plays a large role in many extremal problems in graph theory, the critical number plays a large role in extremal problems on matroids, which are often motivated by analogous problems for graphs. As there is a notion of a graph G containing a copy of a subgraph H , there is a corresponding notion for matroids: a matroid M *contains* a copy of a matroid N if there is a linear injection $\iota : \mathbb{F}_2^{r(N)} \rightarrow \mathbb{F}_2^{r(M)}$ such that $\iota(N) \subseteq M$. We often simply write this as $N \subseteq M$. Many extremal problems pose questions about criteria for the containment or avoidance of a fixed matroid N in a matroid M .

One example of such an extremal problem is to determine the critical threshold of a matroid, another concept inspired by a graph-theoretical analogue.

Definition 1.1. Given a matroid N , the *critical threshold* $\theta(N)$ is the infimum of all $\alpha > 0$ for which there exists $c < \infty$ such that $|M| \geq \alpha 2^{r(M)}$ implies either $N \subseteq M$ or $\chi(M) \leq c$.

The conjecture below is the extremal problem that motivates our work in this paper.

Conjecture 1.2 (Geelen, Nelson, [7, Conj 1.7]). *If $\chi(N) = c$, then $\theta(N) = 1 - i2^{-c}$, where $i \in \{2, 3, 4\}$.*

A more precise and technical version of this conjecture is stated as Conjecture 2.13. The technical details, and previous work towards solving the conjecture, are discussed in Section 2.

The graph-theoretic analog of Conjecture 1.2 is Theorem 2.7, which was proven in [2]. The proof there makes use of the Szemerédi regularity lemma and a corresponding counting lemma. Roughly speaking, the regularity lemma states that, for any desired degree of uniformity ε , the vertices of a sufficiently large graph G can be partitioned into a bounded number of parts of approximately the same size such that most (all but ε -fraction) pairs of parts (X, Y) are ε -uniform, meaning that the edge density $d(X', Y')$ between large enough subsets X', Y' of X, Y does not differ too much from the edge density $d(X, Y)$ between X and Y . Given such a partition Π of G , we can construct a “reduced graph” $R = R_{\varepsilon, \delta}(\Pi)$ whose vertices are the parts in Π , with an edge between a pair (X, Y) if and only if (X, Y) is ε -uniform and $d(X, Y) \geq \delta$. The counting lemma states that for any graph H contained as a subgraph in R , many copies of it are contained in G .

It is natural to consider approaching the critical threshold problem using analogous methods. Various regularity results analogous to the Szemerédi regularity lemma have been shown in the matroid setting, usually framed in terms

of the indicator function for a matroid M decomposing into several parts. The main such result we use, Theorem 3.11, is stated in Section 3.1. The statement of this theorem involves some technical terminology relating to nonclassical polynomial factors and Gowers norms, for which a brief introduction is given in Section 3.1.

In this paper, we build on work in [3] and [10] to develop a corresponding counting lemma for matroids, Theorem 3.13. Again, stating this Counting Lemma in a precise form requires building up technical definitions for concepts like the reduced matroid, based on the results of applying Theorem 3.11. The Counting Lemma and its proof can be found in Section 3.2.

In Section 4 we demonstrate a few simple applications of this counting lemma, giving short new proofs for the matroid analogues of the Removal Lemma and the Erdős-Stone Theorem in graph theory. Finally, we discuss an approach to applying our counting lemma and related techniques to Conjecture 1.2 in Section 5, giving a conditional result on a special case as a demonstration. Along the way, we prove the following technical result, a generalization of the Bose-Burton theorem stated in Section 2 which may be useful in other settings as well.

Proposition 1.3. *Let n, c be positive integers, let k_1, \dots, k_n be nonnegative integers, and let $G = \bigoplus_{i=1}^n \frac{1}{2^{k_i+1}} \mathbb{Z} / \mathbb{Z}$. Let H be a subgroup of G . Let M_1, \dots, M_{2^c-1} be subsets of G . Then there exist $H_1, \dots, H_c \in G/H$, cosets of H , such that for $1 \leq i \leq c$,*

$$\frac{1}{|H|} \sum_{x \in [0,1]^{i-1}} \left| M_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}} \cap \left(H_i + \sum_{j=1}^{i-1} x_j H_j \right) \right| \geq \sum_{j=2^{i-1}}^{2^i-1} \frac{|M_j|}{|G|}. \quad (*)$$

2 Extremal Problems on Graphs and Matroids

We start by looking at a basic extremal problem on graphs, that of avoiding a fixed subgraph H .

Definition 2.1. The *extremal number* for a graph H and integer n is defined by

$$\text{ex}(H, n) = \max\{|E(G)| \mid |G| = n, H \not\subseteq G\}.$$

The following theorem is a classical result.

Theorem 2.2 (Erdős-Stone).

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(H; n)}{|K_n|} = 1 - \frac{1}{\chi(H) - 1}.$$

The special case where $H = K_m$ is a form of *Turán's theorem*.

We can analyze the situation more carefully by looking for density thresholds above which, though graphs G avoiding H may exist, they are constrained by properties like a bounded chromatic number. It turns out that for graphs, the

appropriate notion of density to consider here is the minimum degree $\delta(G)$ of a graph G .

Definition 2.3. Given a graph H , the *chromatic threshold* $\theta(H)$ is the infimum of all $\alpha > 0$ for which there exists $c < \infty$ such that $\delta(G) \geq \alpha|G|$ implies either $H \subseteq G$ or $\chi(G) \leq c$.

The definition of the critical threshold of a matroid was motivated in analogy to Definition 2.3.

The chromatic threshold was first determined for complete graphs in [8], with an explicit sharp bound on the chromatic number involved.

Theorem 2.4 (Goddard, Lyle, [8, Thm 11]). *If $\delta(G) > (2r - 5)n/(2r - 3)$ and $K_r \not\subseteq G$, then $\chi(G) \leq r + 1$. In particular, $\theta(G) \leq \frac{2r-5}{2r-3}$.*

The chromatic threshold of a general graph H was determined in the general case by Allen et al. in [2]. To state the result, we first need to make the following definitions.

Definition 2.5. The *decomposition family* $\mathcal{M}(H)$ of an r -partite graph H is the set of bipartite graphs obtained by deleting all but 2 color classes in some r -coloring of H .

Definition 2.6. A graph H is *r -near-acyclic* if $\chi(H) = r$ and deleting all but 3 color classes in some r -coloring of H yields a graph H' that can be partitioned into a forest F and an independent set S such that every odd cycle in H' meets S in at least 2 vertices.

Now we can state the main result of [2].

Theorem 2.7 (Allen, et al., [2, Thm 2]). *If $\chi(H) = r$, then $\theta(H) = 1 - \frac{1}{r-\frac{i}{2}}$, where $i = 2$ if and only if $\mathcal{M}(H)$ contains no forest, and $i = 4$ if and only if H is r -near-acyclic.*

We can ask the same extremal questions for matroids.

Definition 2.8. The *extremal number* for a matroid N and integer n is defined by

$$\text{ex}(N, n) = \max\{|M| \mid r(M) = n, N \not\subseteq M\}.$$

Note that if $\chi(N) = c$, then N is contained in $BB(n, c)$ for some n . Geelen and Nelson prove the following analogue of the Erdős-Stone theorem in [4].

Theorem 2.9 (Geometric Erdős-Stone, [4, Thm 1.3]).

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(N, n)}{2^n - 1} = 1 - 2^{1-\chi(N)}.$$

The $\chi(N) = 1$ case is known as the *Binary Density Hales-Jewett* theorem. In this case, Geelen and Nelson in fact show that $\text{ex}(AG(k, 2), n) < 2^{\alpha_k n + 1}$, where $\alpha_k = 1 - 2^{-(k-1)}$ [5].

The special case where $N = PG(c - 1, 2)$ is a form of the *Bose-Burton theorem*, which has a more precise statement as follows.

Theorem 2.10. *If M does not contain a copy of $PG(c-1, 2)$, then $|M| \leq 2^{r(M)} - 2^{r(M)-c+1}$.*

Note that taking $G = \mathbb{F}_2^{r(M)}$, H the trivial subgroup, and $M_1 = \dots = M_{2^c-1} = M$ in Proposition 1.3 immediately yields Theorem 2.10.

In [15], Tidor proved a result on the chromatic thresholds of projective geometries, analogous to Goddard and Lyle's result for complete graphs.

Theorem 2.11 (Tidor, [15, Thm 1.4]). *If $|M| > (1 - 3 \cdot 2^{-t})2^{r(M)}$ and $PG(t-1, 2) \not\subseteq M$, then $\chi(M) \in \{t-1, t\}$. In particular, $\theta(M) \leq 1 - 3 \cdot 2^{-t}$.*

To formulate the precise version of Conjecture 1.2, we make the following definition.

Definition 2.12. A matroid M is *c-near-independent* if $\chi(M) = c$ and for some $(c-2)$ -codimensional subspace H with $\chi(M \cap H) = 2$, H has a 1-codimensional subspace S such that $M \cap S$ is linearly independent, and every odd circuit in $M \cap H$ contains at least four elements of $H \setminus S$.

Now we state the precise form of the conjecture.

Conjecture 2.13 (Geelen, Nelson, [7, Conj 5.2]). *If $\chi(N) = c$, then $\theta(N) = 1 - i2^{-c}$, where $i = 2$ if and only if no $(c-1)$ -codimensional subspace S exists such that $S \cap N$ is a set of linearly independent vectors, $i = 4$ if and only if N is *c-near-independent*, and $i = 3$ otherwise.*

In [7], Geelen and Nelson show that the conjectured expression is a valid lower bound.

Theorem 2.14 (Geelen, Nelson, [7, Thm 5.4]). *If $\chi(N) = c$, then $\theta(N) \geq 1 - i2^{-c}$, where $i = 2$ if and only if no $(c-1)$ -codimensional subspace S exists such that $S \cap N$ is a set of linearly independent vectors, and $i = 4$ if and only if N is *c-near-independent*, and $i = 3$ otherwise.*

Combined with the trivial upper bound of $\theta(N) \leq 1 - 2^{1-c}$ that follows immediately from Theorem 2.9, it remains to show that $\theta(N) \leq 1 - 3 \cdot 2^{-c}$ when N has a $(c-1)$ -codimensional flat that is independent, and that $\theta(N) \leq 1 - 4 \cdot 2^{-c}$ when N is *c-near-independent*.

For $\ell \geq c+k-1$, $c > 1$, define $N_{\ell,c,k}$ to be the rank ℓ matroid consisting of the union of $BB(\ell, c-1)$ with k linearly independent vectors contained inside the complement of $BB(\ell, c-1)$ in \mathbb{F}_2^ℓ . This represents the most general maximal case of matroids of critical number c for which i is conjectured to be 3. In Section 5, we verify Conjecture 2.13 for $N_{\ell,2,1}$ and discuss, using a conditional result as a demonstration, how the tools in this paper could be applied to the general $N_{\ell,c,1}$ case.

3 Regularity and Counting

A *regularity* or *decomposition* result, in general, splits a generic object (e.g. a graph, a subset of an abelian group, or a function) into a structured part,

a uniform part, and possibly a small error. A corresponding *counting lemma* then guarantees that the number of copies of a suitable subobject contained in this object can be well-approximated by the number of copies contained in the structured part. The most well-known example of such a pair of results is the Szemerédi regularity lemma and the corresponding counting lemma for subgraphs contained in the reduced graph. As mentioned in the introduction, the use of this pair of lemmas is key to the argument used in [2] to prove Theorem 2.7.

A simple example of an analogous regularity result for matroids is Green’s regularity lemma (specialized to \mathbb{F}_2^n). To state it, we make a few preliminary definitions.

Definition 3.1. Let $V = \mathbb{F}_2^n$. A set $X \subset V$ is *linearly ε -uniform* in V if $|\widehat{1_X}(\xi)| \leq \varepsilon$ for all nonzero $\xi \in \hat{V} = V$, or equivalently if for each hyperplane $H \leq V$,

$$||X \cap H| - |X \setminus H|| \leq \varepsilon|V|.$$

Definition 3.2. Let $X \subseteq V = \mathbb{F}_2^n$. A subspace $W \leq V$ is *linearly ε -regular* with respect to X if for all but $\varepsilon|V|$ values of $v \in V$, $X - v \cap W$ is linearly ε -uniform in W .

Green’s regularity result is the following.

Theorem 3.3 (Geometric Regularity Lemma, [9, Thm 2.1]). *For any $\varepsilon \in (0, \frac{1}{2})$ there is a $T > 0$ such that for any $V = \mathbb{F}_2^n$ and any subset $X \subset V$ there is a subspace $W \subseteq V$ of codimension at most T that is linearly ε -regular with respect to X .*

This notion of regularity readily yields a counting lemma for triangles (and indeed, all odd circuits) in matroids, which Geelen and Nelson use in their proofs that $\theta(PG(1, 2)) \leq \frac{1}{4}$ [7] and that odd circuits have critical threshold 0 [6]. In Section 5.1, we use a method along the same lines as their proof to verify Conjecture 2.13 for $N_{\ell, 2, 1}$.

Unfortunately, the linear Fourier-analytic notion of regularity provided by Theorem 3.3 is not strong enough for a counting lemma to hold for general submatroids N . In attempting to translate the ideas of [2] into tools for the matroid threshold problem, we therefore need a stronger regularity statement, one that admits a corresponding, more general counting lemma.

3.1 Regularity on Matroids

After the inverse conjecture for the Gowers norm over finite fields of low characteristic was established [14], stronger regularity results, using regularity with respect to the Gowers norms, came within reach. The primary regularity result that we will use is stated in [3] as a decomposition theorem for bounded functions on \mathbb{F}^n . To work with this result, we will first need to introduce a few technical concepts from higher-order Fourier analysis.

We will only be concerned with the field $\mathbb{F} = \mathbb{F}_2$, though most of the concepts below also extend to prime-ordered fields \mathbb{F}_p in general. Given a function $f : \mathbb{F}^n \rightarrow \{0, 1\}$, in our case usually the indicator function of some matroid $M \subseteq \mathbb{F}^n \setminus \{0\}$, the decomposition theorem will split it into a sum of three parts: a structured part, a uniform part, and a small error. Here we will address the technical issues that arise in working with the first two parts. In the sections below, we largely quote the terminology and notation used in [3] and [10].

3.1.1 The Gowers norm and nonclassical polynomials

Definition 3.4. Given a function $f : \mathbb{F}^n \rightarrow \mathbb{C}$ and an integer $d \geq 1$, the *Gowers norm of order d* for f is

$$\|f\|_{U^d} = \left| \mathbf{E}_{h_1, \dots, h_d, x \in \mathbb{F}^n} \left[\prod_{i_1, \dots, i_d \in \{0, 1\}} \mathcal{C}^{i_1 + \dots + i_d} f \left(x + \sum_{j=1}^d i_j h_j \right) \right] \right|^{1/2^d},$$

where \mathcal{C} denotes the conjugation operator.

It is easy to see that $\|f\|_{U^d}$ is increasing in d and is indeed a norm for $d \geq 2$, and that $\|f\|_{U^1} = |\mathbf{E}[f]|$ and $\|f\|_{U^2} = \|\hat{f}\|_{l^4}$. So, the Gowers norm of order 2 is related to the Fourier bias used in Green's regularity lemma, a measure of correlation with exponentials of linear polynomials: $\|f\|_{U^2}$ is large if and only if $\sup_{\xi \neq 0} |\hat{f}(\xi)|$ is large, i.e. if and only if f is strongly correlated with the exponential of some linear polynomial. It is natural to expect the Gowers norm of order $d + 1$ to be similarly related to polynomials of degree d ; conjectures that a large Gowers- $(d + 1)$ norm implies correlation with the exponential of a degree d polynomial, in various settings, were known as *inverse conjectures* for the Gowers norms.

For large d over fields of small characteristic, it turns out that the inverse conjectures are not true as stated; the right notion to consider, over which an inverse theorem for the Gowers norms actually holds, is that of a *nonclassical polynomial*.

Definition 3.5. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. Given an integer $d \geq 0$, a function $P : \mathbb{F}^n \rightarrow \mathbb{T}$ is called a *(non-classical) polynomial of degree at most d* if for all $h_1, \dots, h_d, x \in \mathbb{F}^n$,

$$\sum_{i_1, \dots, i_d \in \{0, 1\}} (-1)^{i_1 + \dots + i_d} P \left(x + \sum_{j=1}^d i_j h_j \right) = 0.$$

Since we will be working mostly with non-classical polynomials, it should be assumed that any use of the word “polynomial” refers to a possibly non-classical polynomial unless otherwise specified.

Let $\mathbf{e}(x) = e^{2\pi i x}$. It follows from definition that $\|f\|_{U^{d+1}} = 1$ if and only if $f = \mathbf{e}(P)$ for some non-classical polynomial P of degree at most d , and $\|f \cdot \mathbf{e}(P)\|_{U^{d+1}} = \|f\|_{U^{d+1}}$ for any function f and non-classical polynomial P of

degree at most d . Non-classical polynomials can be characterized in terms of classical ones by the following lemma of Tao and Ziegler [14].

Lemma 3.6 ([14, Lemma 1.7]). *Let $|\cdot|$ denote the standard map from \mathbb{F}_p to $\{0, 1, \dots, p-1\}$. A function $P : \mathbb{F}_p^n \rightarrow \mathbb{T}$ is a polynomial of degree at most d if and only if P can be represented as*

$$P(x_1, \dots, x_n) = \alpha + \sum_{\substack{0 \leq d_1, \dots, d_n < p; k \geq 0: \\ 0 < \sum_i d_i \leq d-k(p-1)}} \frac{c_{d_1, \dots, d_n, k} |x_1|^{d_1} \cdots |x_n|^{d_n}}{p^{k+1}} \mod 1,$$

for a unique choice of $c_{d_1, \dots, d_n, k} \in \{0, 1, \dots, p-1\}$ and $\alpha \in \mathbb{T}$. The element α is called the *shift* of P , and the largest integer k such that there exist d_1, \dots, d_n for which $c_{d_1, \dots, d_n, k} \neq 0$ is called the *depth* of P . A *depth- k polynomial* P takes values in a coset of the subgroup $\mathbb{U}_{k+1} := \frac{1}{p^{k+1}}\mathbb{Z}/\mathbb{Z}$. Classical polynomials correspond to polynomials with 0 shift and 0 depth.

For convenience we will assume henceforth that all polynomials have shift 0, so that all polynomials of depth k take values in \mathbb{U}_{k+1} ; this will not affect our arguments.

3.1.2 Polynomial factors and rank

Definition 3.7. A *polynomial factor* \mathcal{B} of \mathbb{F}^n is a partition of \mathbb{F}^n into finitely many pieces, called *atoms*, such that for some polynomials P_1, \dots, P_C , each atom is defined as the solution set $\{x | \forall i \in \{1, \dots, C\} P_i(x) = b_i\}$ for some $(b_1, \dots, b_C) \in \mathbb{T}^C$. The *complexity* of \mathcal{B} is the number of defining polynomials $|\mathcal{B}| = C$, and the *degree* is the highest degree among P_1, \dots, P_C . If P_i has depth k_i , the *order* of \mathcal{B} is $\|\mathcal{B}\| = \prod_{i=1}^C p^{k_i+1}$, an upper bound on the number of atoms in \mathcal{B} .

Definition 3.8. The *d -rank* $\text{rank}_d(P)$ of a polynomial P is the smallest integer r such that P can be expressed as a function of r polynomials of degree at most $d-1$. The *rank* of a polynomial factor defined by P_1, \dots, P_C is the least integer r for which there is a tuple $(\lambda_1, \dots, \lambda_C) \in \mathbb{Z}^C$, with $(\lambda_1 \mod p^{k_1+1}, \dots, \lambda_C \mod p^{k_C+1}) \neq 0^C$, such that $\text{rank}_d(\sum_{i=1}^C \lambda_i P_i) \leq r$, where $d = \max_i \deg(\lambda_i P_i)$.

Given a polynomial factor \mathcal{B} and a function $r : \mathbb{Z}_{>0} \rightarrow \mathbb{Z}_{>0}$, we say that \mathcal{B} is *r -regular* if \mathcal{B} is of rank larger than $r(|\mathcal{B}|)$.

We can also define an analytic notion of uniformity for polynomial factors.

Definition 3.9. For $\varepsilon > 0$, we say that \mathcal{B} is *ε -uniform* if for all $(\lambda_1, \dots, \lambda_C) \in \mathbb{Z}^C$ with $(\lambda_1 \mod p^{k_1+1}, \dots, \lambda_C \mod p^{k_C+1}) \neq 0^C$,

$$\left\| \mathbf{e} \left(\sum_i \lambda_i P_i \right) \right\|_{U^d} < \varepsilon.$$

The following fact, noted in [10], follows from Theorem 1.20 of [14].

Proposition 3.10. *For every $\varepsilon > 0$, $d \in \mathbb{Z}_{>0}$ there exists an integer $r = r(d, \varepsilon)$ such that every r -regular degree d polynomial factor \mathcal{B} is ε -uniform.*

3.1.3 The Strong Decomposition Theorem

We can finally state the main regularity result we will use, the strong decomposition theorem of [3].

Theorem 3.11 (Strong Decomposition Theorem, [3, Theorem 5.1]). *Suppose $\delta > 0$ and $d \geq 1$ are integers. Let $\eta : \mathbb{N} \rightarrow \mathbb{R}^+$ be an arbitrary non-increasing function and $r : \mathbb{N} \rightarrow \mathbb{N}$ be an arbitrary non-decreasing function. Then there exist $N = N(\delta, \eta, r, d)$ and $C = C(\delta, \eta, r, d)$ such that the following holds.*

Given $f : \mathbb{F}^n \rightarrow \{0, 1\}$ where $n > N$, there exist three functions $f_1, f_2, f_3 : \mathbb{F}^n \rightarrow \mathbb{R}$ and a polynomial factor \mathcal{B} of degree at most d and complexity at most C such that the following conditions hold:

- (i) $f = f_1 + f_2 + f_3$.
- (ii) $f_1 = \mathbf{E}[f|\mathcal{B}]$, the expected value of f on an atom of \mathcal{B} .
- (iii) $\|f_2\|_{U^{d+1}} \leq \eta(|\mathcal{B}|)$.
- (iv) $\|f_3\|_2 \leq \delta$.
- (v) f_1 and $f_1 + f_3$ have range $[0, 1]$; f_2 and f_3 have range $[-1, 1]$.
- (vi) \mathcal{B} is r -regular.

In analogy with the terminology for the Szemerédi regularity lemma, we will call a decomposition of $f = 1_M$ with the properties given by Theorem 3.11 for parameters δ, η, r, d a (δ, η, r, d) -regular partition of f (or of M), and we say that \mathcal{B} is its corresponding factor. Similarly, an (η, r, d) -regular partition of f (or of M) is the same thing with an unspecified value for δ .

Theorem 3.11 is strong enough to help us prove a corresponding counting lemma for general binary matroids.

3.2 The Counting Lemma

Before stating our Counting Lemma, we first we define the notion of a *reduced matroid*, in analogy to the reduced graph used in the graph counting lemma.

Definition 3.12 (Reduced Matroid). Given a matroid $M \subseteq \mathbb{F}^n \setminus \{0\}$ and an (η, r, d) -regular partition $f_1 + f_2 + f_3$ of M with corresponding factor \mathcal{B} , for any $\varepsilon, \zeta > 0$ define the (ε, ζ) -reduced matroid $R = R_{\varepsilon, \zeta}$ to be the subset of \mathbb{F}^n whose indicator function F is constant on each atom b of \mathcal{B} and equals 1 if and only if

1. $\mathbf{E}[|f_3(x)|^2 \mid x \in b] \leq \varepsilon^2$, and
2. $\mathbf{E}[f(x) \mid x \in b] \geq \zeta$.

So, R gives the atoms of the decomposition in which M has high density and the L^2 error term is small.

As in the counting lemma for graphs, it will turn out that we do not need the hypothesis of having a *copy* of N in $R_{\varepsilon, \zeta}$, i.e. an *injective* linear map sending N inside $R_{\varepsilon, \zeta}$. Rather, we only need there to be a *homomorphism* from N to $R_{\varepsilon, \zeta}$, i.e. any linear map ι such that $\iota(N) \subseteq R_{\varepsilon, \zeta}$.

Theorem 3.13 (Counting Lemma). *For every matroid N , positive real number ζ , and integer $d \geq |N| - 2$, there exist positive real numbers β and ε_0 , a positive nonincreasing function $\eta : \mathbb{Z}^+ \rightarrow \mathbb{R}^+$, and positive nondecreasing functions $r, \nu : \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ such that the following holds for all $\varepsilon \leq \varepsilon_0$. Let $M \subseteq \mathbb{F}^n \setminus \{0\}$ be a matroid with an (η, r, d) -regular partition $f_1 + f_2 + f_3$ with corresponding factor \mathcal{B} . If $n \geq \nu(|\mathcal{B}|)$ and there exists a homomorphism from N to the reduced matroid $R_{\varepsilon, \zeta}$, then there exist at least $\beta \frac{(2^n)^{r(N)}}{|\mathcal{B}|^{|N|}}$ copies of N in M .*

The case where N is an affine matroid (i.e. $\chi(N) = 1$) is proved in [3] in the context of property testing; here we prove the lemma in full generality using a very similar argument. The basic idea is to obtain a lower bound for the probability that a linear map $\iota : \mathbb{F}^{r(N)} \rightarrow \mathbb{F}^n$ chosen uniformly at random sends N to a set contained entirely within M , by splitting $f = 1_M$ into its three parts and expanding out the product we get in the expectation expression. Letting $N = \{N_1, \dots, N_m\}$ and $r(N) = \ell$ we have

$$\Pr_{\iota: \mathbb{F}^\ell \rightarrow \mathbb{F}^n} [\iota(N) \subseteq M] = \mathbf{E}_\iota \left[\prod_{i=1}^m f(\iota(N_i)) \right] = \mathbf{E}_\iota \left[\sum_{(j_1, \dots, j_m) \in [1, 3]^m} \prod_{i=1}^m f_{j_i}(\iota(N_i)) \right].$$

For ease of notation in the proof that follows, we will introduce the concept of a linear form, as used in [3] and [10].

Definition 3.14. A *linear form* on k variables is a linear map $L : (\mathbb{F}^n)^k \rightarrow \mathbb{F}^n$. If it is given by $L(x_1, \dots, x_k) = \sum_{i=1}^k \ell_i x_i$, where $\ell_i \in \mathbb{F} \ \forall i$, we write $L = (\ell_1, \dots, \ell_k)$.

Note that a linear form on k variables can be thought of as a vector in \mathbb{F}^k . Conversely, we can think of each element N_j of N as a linear form L_j on ℓ variables. Each linear map $\iota : \mathbb{F}^\ell \rightarrow \mathbb{F}^n$ corresponds to a point $X = (x_1, \dots, x_\ell) \in (\mathbb{F}^n)^\ell$ such that $\iota(N_j) = L_j(X)$. So, instead of taking an expectation over linear maps, we can take the more intuitive approach of taking an expectation over tuples of points. The expression from before is the same as

$$\begin{aligned} \Pr_{X \in (\mathbb{F}^n)^\ell} [L_j(X) \in M \ \forall j \in [1, m]] &= \mathbf{E}_X \left[\prod_{i=1}^m f(L_j(X)) \right] \\ &= \mathbf{E}_X \left[\sum_{(i_1, \dots, i_m) \in [1, 3]^m} \prod_{i=1}^m f_{i_j}(L_j(X)) \right]. \end{aligned}$$

The terms involving the Gowers uniform part f_2 are the easiest to deal with; we will simply invoke the following result.

Lemma 3.15 ([10, Lemma 3.12]). *Let $f_1, \dots, f_m : \mathbb{F}^n \rightarrow \mathbb{D}$. Let $N = \{L_1, \dots, L_m\}$ be a system of linear forms in ℓ variables. Then for $s \geq m - 2$,*

$$\left| \mathbf{E}_{X \in (\mathbb{F}^n)^\ell} \left[\prod_{j=1}^m f_j(L_j(X)) \right] \right| \leq \min_{1 \leq j \leq m} \|f_j\|_{U^{s+1}}.$$

To deal with the remaining terms, we will use a near-orthogonality theorem from [10]. To state this near-orthogonality theorem, we first need to introduce the notion of *consistency* as defined in [10]. By a *homogeneous* polynomial over \mathbb{F}_p we mean a polynomial P such that for all $c \in \mathbb{F}_p$ there exists a $c' \in \mathbb{F}_p$ such that $P(cx) \equiv c'P(x)$. In the case of $\mathbb{F} = \mathbb{F}_2$, this restriction is equivalent to simply requiring $P(0) = 0$.

Definition 3.16 (Consistency). Let $N = \{L_1, \dots, L_m\}$ be a system of linear forms in ℓ variables. A vector $(\beta_1, \dots, \beta_m) \in \mathbb{T}^m$ is said to be (d, k) -consistent with N if there exists a homogeneous polynomial P of degree d and depth k and a point $X \in (\mathbb{F}^n)^\ell$ such that $P(X(L_j)) = \beta_j$ for every $j \in [m]$. Let $\Phi_{d,k}(N)$ denote the set of all such vectors. This is a subgroup of \mathbb{U}_{k+1}^m ; we define

$$\Phi_{d,k}(N)^\perp := \left\{ (\lambda_1, \dots, \lambda_m) \in [0, 2^{k+1} - 1]^m : \forall (\beta_1, \dots, \beta_m) \in \Phi_{d,k}(N), \sum \lambda_j \beta_j = 0 \right\},$$

the set of all $(\lambda_1, \dots, \lambda_m) \in [0, 2^{k+1} - 1]^m$ such that $\sum_{k=1}^m \lambda_j P(L_j(X)) \equiv 0$ for every homogeneous polynomial P of degree d and depth k and every point X . Call $\Phi_{d,k}(N)^\perp$ the (d, k) -dependency set of N .

Theorem 3.17 (Near Orthogonality, [10, Theorem 3.7]). *Let $N = \{L_1, \dots, L_m\}$ be a system of linear forms in ℓ variables, and let $\mathcal{B} = (P_1, \dots, P_C)$ be an ε -uniform polynomial factor for some $\varepsilon \in (0, 1]$ defined only by homogeneous polynomials. For every tuple Λ of integers $(\lambda_{i,j})_{i \in [C], j \in [m]}$, define*

$$P_\Lambda(X) = \sum_{i \in [C], j \in [m]} \lambda_{i,j} P_i(L_j(X)).$$

Then one of the following two statements holds:

- $P_\Lambda \equiv 0$, and furthermore for every $i \in [C]$, we have $(\lambda_{i,j})_{j \in [m]} \in \Phi_{d_i, k_i}(N)^\perp$, where d_i, k_i are the degree and depth of P_i , respectively.
- P_Λ is non-constant and $|\mathbf{E}_{X \in (\mathbb{F}^n)^\ell} [e(P_\Lambda)]| < \varepsilon$.

Remark. Over a general prime-ordered field \mathbb{F}_p , the restriction of homogeneity is dealt with by modifying the decomposition theorem to only use homogeneous polynomials in the factor \mathcal{B} . The situation is even simpler over $\mathbb{F} = \mathbb{F}_2$, where any polynomial factor can be rewritten in terms of homogeneous polynomials simply by shifting each polynomial by a constant.

Directly applying Theorem 3.17 yields the following result, which estimates the number of copies of N with each element in a specified atom of \mathcal{B} .

Theorem 3.18 (Near-equidistribution, [10, Theorem 3.10]). *Given $\varepsilon > 0$, let \mathcal{B} be an ε -uniform polynomial factor of degree $d > 0$ and complexity C that is defined by a tuple of homogeneous polynomials $P_1, \dots, P_C : \mathbb{F}^n \rightarrow \mathbb{T}$ having respective degrees d_1, \dots, d_C and depths k_1, \dots, k_C . Let $N = \{L_1, \dots, L_m\}$ be a system of linear forms on ℓ variables.*

Suppose $(\beta_{i,j})_{i \in [C], j \in [m]} \in \mathbb{T}^{C \times m}$ is such that $(\beta_{i,1}, \dots, \beta_{i,m}) \in \Phi_{d_i, k_i}(N)$ for every $i \in [C]$. Then

$$\left| \Pr_{X \in (\mathbb{F}^n)^\ell} [P_i(L_j(X)) = \beta_{i,j} \ \forall i \in [C], j \in [m]] - \frac{1}{K} \right| \leq \varepsilon,$$

where $K = \prod_{i=1}^C |\Phi_{d_i, k_i}(N)|$.

In particular, taking N to have a single element yields an estimate on the size of each atom.

Corollary 3.19 (Size of atoms, [3, Lemma 3.2]). *Given $\varepsilon > 0$, let \mathcal{B} be an ε -uniform polynomial factor of degree $d > 0$ and complexity C that is defined by a tuple of homogeneous polynomials $P_1, \dots, P_C : \mathbb{F}^n \rightarrow \mathbb{T}$ having respective depths k_1, \dots, k_C .*

Suppose $b = (b_1, \dots, b_C) \in \mathbb{T}^C$ is such that $b_i \in \mathbb{U}_{k_i+1}$ for every $i \in [C]$. Then

$$\left| \Pr_x [P_i(x) = b_i \ \forall i \in [C]] - \frac{1}{\|\mathcal{B}\|} \right| \leq \varepsilon.$$

Now we proceed with the proof of the Counting Lemma.

Proof of Theorem 3.13. Let $\ell = r(N)$ and $\alpha(C) = 2^{-2dCm}$. We set $r(C)$ to be the integer $r(d, \alpha(C))$ given by Proposition 3.10 such that every r -regular degree d polynomial factor is $\alpha(C)$ -uniform. Let $\mathcal{B} = (P_1, \dots, P_C)$. So, \mathcal{B} is $\alpha(C)$ -uniform. Notice that $\|\mathcal{B}\| = \prod_{i=1}^C 2^{k_i+1} \leq 2^{dC}$.

We want a lower bound for the probability that for a linear map $\iota : \mathbb{F}^\ell \rightarrow \mathbb{F}^n$ chosen uniformly at random, each element of N is sent inside M . Since degenerate maps (ones where the image of N is of lower rank than N) are sparse, this will give us that a constant fraction of the copies of N in \mathbb{F}^n (depending on $\|\mathcal{B}\|$ in an appropriate way) are contained in M .

As before, we represent N as a system $\{L_1, \dots, L_m\}$ of linear forms on ℓ variables. The probability we want to bound is then

$$\begin{aligned} \Pr_{X \in (\mathbb{F}^n)^\ell} [L_j(X) \subseteq M \ \forall j \in [1, m]] &= \mathbf{E}_X \left[\prod_{j=1}^m f(L_j(X)) \right] \\ &= \sum_{(i_1, \dots, i_m) \in [1, 3]^m} \mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) \right]. \end{aligned}$$

There are 3^m terms in the sum. One of these, the one that will turn out to be the main term, involves only f_1 . Of the rest, $2^m - 1$ involve f_1 and f_3 but not f_2 , and the other $3^m - 2^m$ terms involve f_2 .

If one of the i_j is 2, then since $d \geq m - 2$, by Lemma 3.15 we have

$$\left| \mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) \right] \right| \leq \min_{1 \leq j \leq m} \|f_{i_j}\|_{U^{d+1}} \leq \|f_2\|_{U^{d+1}} \leq \eta(|\mathcal{B}|).$$

We can choose η later to make all of these terms sufficiently small. Our probability is thus at least

$$\sum_{(i_1, \dots, i_m) \in \{1, 3\}^m} \mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) \right] - 3^m \eta(|\mathcal{B}|).$$

To deal with the remaining terms, we restrict to counting within certain “good” atoms of the decomposition. Specifically, let Y be a point such that $L_1(Y), \dots, L_m(Y) \in R_{\varepsilon, \zeta}$, i.e. $(L_1(Y), \dots, L_m(Y))$ forms a copy of the matroid N in $R_{\varepsilon, \zeta}$. Let $\beta_{i,j} = P_i(L_j(Y))$, and let $b_j = (\beta_{1,j}, \dots, \beta_{C,j})$ be the atom of \mathcal{B} that $L_j(Y)$ is in. We restrict to counting across tuples X such that $L_j(X) \in b_j$ for all j . Since $f_1 + f_3$ is always nonnegative,

$$\begin{aligned} & \sum_{(i_1, \dots, i_m) \in \{1, 3\}^m} \mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) \right] \\ & \geq \sum_{(i_1, \dots, i_m) \in \{1, 3\}^m} \mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) 1_{[\mathcal{B}(L_j(X)) = b_j]} \right]. \end{aligned}$$

We next deal with the main term. Since $f_1 \geq 0$, applying Theorem 3.18 with $\varepsilon = \alpha(C)$ gives

$$\begin{aligned} & \mathbf{E}_X \left[\prod_{j=1}^m f_1(L_j(X)) 1_{[\mathcal{B}(L_j(X)) = b_j]} \right] \\ & = \Pr_{X \in (\mathbb{R}^n)^\ell} [P_i(L_j(X)) = \beta_{i,j} \ \forall i \in [C], j \in [m]] \cdot \\ & \quad \mathbf{E}_X \left[\prod_{j=1}^m f_1(L_j(X)) \middle| \forall j \in [m], \mathcal{B}(L_j(X)) = b_j \right] \\ & \geq \left(\frac{1}{K} - \alpha(C) \right) \zeta^m. \end{aligned}$$

Now we deal with the terms involving f_3 , following the argument used in the corresponding part of the proof of Theorem 5.10 in [3].

Take such a term $\mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) 1_{[\mathcal{B}(L_j(X))=b_j]} \right]$, and suppose $i_k = 3$. Without loss of generality we can take a linear transformation of coordinates and assume $L_k(x_1, \dots, x_\ell) = x_1$. We can also assume without loss of generality that $k = 1$. Then, since $|f_1|, |f_3| \leq 1$, we have

$$\begin{aligned} & \left| \mathbf{E}_X \left[\prod_{j=1}^m f_{i_j}(L_j(X)) 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \right| \leq \mathbf{E}_X \left[|f_3(x_1)| \prod_{j=1}^m 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \\ &= \mathbf{E}_{x_1} \left[|f_3(x_1)| 1_{[\mathcal{B}(x_1)=b_1]} \mathbf{E}_{x_2, \dots, x_\ell} \left[\prod_{j=1}^m 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \right]. \end{aligned}$$

By Cauchy-Schwarz,

$$\begin{aligned} & \left(\mathbf{E}_X \left[|f_3(x_1)| \prod_{j=1}^m 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \right)^2 \\ & \leq \mathbf{E}_{x_1} \left[|f_3(x_1)|^2 1_{[\mathcal{B}(x_1)=b_1]} \right] \mathbf{E}_{x_1} \left(\mathbf{E}_{x_2, \dots, x_\ell} \left[\prod_{j=1}^m 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \right)^2. \quad (*) \end{aligned}$$

By Corollary 3.19, $\Pr_{x_1}[\mathcal{B}(x_1) = b_1] \leq \frac{1}{\|\mathcal{B}\|} + \alpha(C)$. Thus by condition (1) in the definition of the reduced matroid,

$$\begin{aligned} & \mathbf{E}_{x_1} \left[|f_3(x_1)|^2 1_{[\mathcal{B}(x_1)=b_1]} \right] = \mathbf{E}_{x_1} \left[|f_3(x_1)|^2 \mid x_1 \in b_1 \right] \Pr_{x_1}[\mathcal{B}(x_1) = b_1] \\ & \leq \varepsilon^2 \left(\frac{1}{\|\mathcal{B}\|} + \alpha(C) \right) \leq \frac{2\varepsilon^2}{\|\mathcal{B}\|}. \end{aligned}$$

Let $Y = (y_2, \dots, y_\ell) \in (\mathbb{F}^n)^{\ell-1}$, so that (x_1, Y) forms another input to the linear forms L_j such that $L_k(x_1, Y) = L_k(X) = x_1$. The second term in the right hand side of (*) expands as

$$\begin{aligned}
& \mathbf{E}_{x_1} \left(\mathbf{E}_{x_2, \dots, x_l} \left[\prod_{j=1}^m 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \right)^2 = \mathbf{E}_{x_1} \left(\mathbf{E}_{x_2, \dots, x_l} \left[\prod_{\substack{i \in [C] \\ j \in [m]}} 1_{[P_i(L_j(X))=\beta_{i,j}]} \right] \right)^2 \\
&= \mathbf{E}_{x_1} \left(\mathbf{E}_{x_2, \dots, x_l} \left[\prod_{\substack{i \in [C] \\ j \in [m]}} \frac{1}{2^{k_i+1}} \sum_{\lambda_{i,j}=0}^{2^{k_i+1}-1} e(\lambda_{i,j}(P_i(L_j(X)) - \beta_{i,j})) \right] \right)^2 \\
&= \frac{1}{\|\mathcal{B}\|^{2m}} \mathbf{E}_{x_1} \left(\sum_{\substack{(\lambda_{i,j}) \in \\ \prod_{i,j} [0, 2^{k_i+1}-1]}} e \left(- \sum_{\substack{i \in [C] \\ j \in [m]}} \lambda_{i,j} \beta_{i,j} \right) \mathbf{E}_{x_2, \dots, x_l} e \left(\sum_{\substack{i \in [C] \\ j \in [m]}} \lambda_{i,j} P_i(L_j(X)) \right) \right)^2 \\
&\leq \frac{1}{\|\mathcal{B}\|^{2m}} \sum_{\substack{(\lambda_{i,j}), (\tau_{i,j}) \in \\ \prod_{i,j} [0, 2^{k_i+1}-1]}} \left| \mathbf{E}_{X,Y} \left[e \left(\sum_{\substack{i \in [C] \\ j \in [m]}} \lambda_{i,j} P_i(L_j(X)) \right) e \left(\sum_{\substack{i \in [C] \\ j \in [m]}} \tau_{i,j} P_i(L_j(x_1, Y)) \right) \right] \right|.
\end{aligned}$$

To bound this expression, we will interpret it as an expectation of the form for which Theorem 3.17 gives upper bounds, for some system of linear forms N' that we construct. Specifically, let N' be a system of linear forms $\tilde{L}_1, \dots, \tilde{L}_{2m-1}$ on $2\ell - 1$ variables, where $\tilde{L}_i(X, Y) = L_i(X)$ for $i \in \{1, \dots, m\}$ and $\tilde{L}_i(X, Y) = L_{i-m+1}(x_1, Y)$ for $i \in \{m+1, \dots, 2m-1\}$. That is, N' corresponds to a rank- $(2\ell - 1)$ matroid consisting of the union of two copies of N with $L_1 = N_1$ as the only shared element.

Lemma 3.20. *For each i , $|\Phi_{d_i, k_i}(N')^\perp| = |\Phi_{d_i, k_i}(N)^\perp|^2$.*

Proof. The proof is essentially the same as that of Lemma 5.13 in [3], except that a key step now uses relevant polynomials being homogeneous instead of relevant linear forms being affine.

Consider the map $\varphi : \Phi_{d_i, k_i}(N)^\perp \times \Phi_{d_i, k_i}(N)^\perp \rightarrow \Phi_{d_i, k_i}(N')^\perp$ given by

$$\varphi((\lambda_1, \dots, \lambda_m), (\tau_1, \dots, \tau_m)) = (\lambda_1 + \tau_1, \lambda_2, \dots, \lambda_m, \tau_2, \dots, \tau_m).$$

If $\lambda = (\lambda_1, \dots, \lambda_m), \tau = (\tau_1, \dots, \tau_m) \in \Phi_{d_i, k_i}(N)^\perp$, then $\sum_{j=1}^m \lambda_j P(L_j(X)) = \sum_{j=1}^m \tau_j P(L_j(X)) = 0$ for all P and X . So, for all (X, Y) ,

$$\sum_{j=1}^{2m-1} \varphi(\lambda, \tau)_j P(\tilde{L}_j(X, Y)) = \sum_{j=1}^m \lambda_j P(L_j(X)) + \sum_{j=1}^m \tau_j P(L_j(x_1, Y)) = 0.$$

Thus φ indeed maps $\Phi_{d_i, k_i}(N)^\perp \times \Phi_{d_i, k_i}(N)^\perp$ to $\Phi_{d_i, k_i}(N')^\perp$. We claim that φ is a bijection, which then implies the desired equality.

Suppose $\lambda \in \Phi_{d_i, k_i}(N)^\perp$. By the definition of a linear form, for each i either $L_i(x_1, 0, \dots, 0) \equiv 0$ or $L_i(x_1, 0, \dots, 0) \equiv x_1$. Let S be the set of i such that $L_i(x_1, 0, \dots, 0) \equiv x_1$. Note that $1 \in S$. Setting $x_2 = \dots = x_\ell = 0$ and setting P, x_1 such that P is a linear polynomial with $P(x_1) = 1, P(0) = 0$ gives

$$0 = \left(\sum_{j \in S} \lambda_j \right) P(x_1) + \left(\sum_{j \notin S} \lambda_j \right) P(0) = \left(\sum_{j \in S} \lambda_j \right) x_1,$$

so $\sum_{j \in S} \lambda_j = 0$, meaning $\lambda_1 = -\sum_{j \in S, j \neq 1} \lambda_j$. Thus λ is uniquely determined given $\lambda_2, \dots, \lambda_m$, and so both λ and τ are uniquely determined given $\varphi(\lambda, \tau)$, meaning φ is injective.

Now suppose $(\lambda_1, \dots, \lambda_m, \tau_2, \dots, \tau_m) \in \Phi_{d_i, k_i}(N')^\perp$, so $\sum_{j=1}^m (\lambda_j Q(L_j(X)) + \sum_{j=2}^m \tau_j Q(x_1, Y)) = 0$ for every Q and (X, Y) . For convenience, let $\tau_1 = \lambda_1$. Similarly to before, setting $x_2 = \dots = x_\ell = 0$ gives

$$\sum_{j \in S} \lambda_j Q(x_1) + \sum_{j=2}^m \tau_j Q(L_j(x_1, Y)) = 0,$$

for any x_1, Y and any homogeneous Q , while setting $y_2 = \dots = y_\ell = 0$ gives

$$\sum_{j=2}^m \lambda_j Q(L_j(X)) + \sum_{j \in S} \tau_j Q(x_1) = 0,$$

for any X and any homogeneous Q . Here we have used the fact that $Q(0) = 0$.

In particular, setting $x_2 = \dots = x_\ell = y_2 = \dots = y_\ell = 0$ and setting Q, x_1 such that Q is a linear polynomial with $Q(x_1) = 1, Q(0) = 0$ gives $0 = \sum_{j \in S} \lambda_j + \sum_{j \in S} \tau_j - \lambda_1 = \sum_{j \in S, j \neq 1} (\lambda_j + \tau_j) - \lambda_1$. So, fixing X ,

$$\begin{aligned} 0 &= \sum_{j=2}^m \lambda_j Q(L_j(X)) + \sum_{j \in S} \tau_j Q(x_1) = \sum_{j=2}^m \lambda_j Q(L_j(X)) - \sum_{j \in S} \lambda_j Q(x_1) + \lambda_1 Q(x_1) \\ &= \left(-\sum_{\substack{j \in S \\ j \neq 1}} \lambda_j \right) Q(x_1) + \sum_{j=2}^m \lambda_j Q(L_j(X)). \end{aligned}$$

Thus $\left(-\sum_{\substack{j \in S \\ j \neq 1}} \lambda_j, \lambda_2, \dots, \lambda_m \right) \in \Phi_{d_i, k_i}(N)^\perp$, and $\left(-\sum_{\substack{j \in S \\ j \neq 1}} \tau_j, \tau_2, \dots, \tau_m \right) \in \Phi_{d_i, k_i}(N)^\perp$ likewise. But

$$\varphi \left(\left(-\sum_{\substack{j \in S \\ j \neq 1}} \lambda_j, \lambda_2, \dots, \lambda_m \right), \left(-\sum_{\substack{j \in S \\ j \neq 1}} \tau_j, \tau_2, \dots, \tau_m \right) \right) = (\lambda_1, \dots, \lambda_m, \tau_2, \dots, \tau_m).$$

So φ is surjective, and thus bijective as desired. \square

Now, let $\mu_i = \varphi(\lambda_i, \tau_i)$. We have

$$\begin{aligned}
& \sum_{\substack{(\lambda_{i,j}), (\tau_{i,j}) \in \\ \prod_{i,j} [0, 2^{k_i+1}-1]}} \left| \mathbf{E}_{X,Y} \left[e \left(\sum_{\substack{i \in [C] \\ j \in [m]}} \lambda_{i,j} P_i(L_j(X)) \right) e \left(\sum_{\substack{i \in [C] \\ j \in [m]}} \tau_{i,j} P_i(L_j(x_1, Y)) \right) \right] \right| \\
&= \left(\prod_{i=1}^C 2^{k_i+1} \right) \sum_{\substack{(\mu_{i,j}) \in \\ \prod_{i \in [C], j \in [2m-1]} [0, 2^{k_i+1}-1]}} \left| \mathbf{E}_{X,Y} \left[e \left(\sum_{\substack{i \in [C] \\ j \in [2m-1]}} \mu_{i,j} P_i(\tilde{L}_j(X, Y)) \right) \right] \right| \\
&\leq \left(\prod_{i=1}^C 2^{k_i+1} \right)^{2m} \alpha(C) + \left(\prod_{i=1}^C 2^{k_i+1} \right) \prod_{i=1}^C |\Phi_{d_i, k_i}(N')^\perp| \\
&= \|\mathcal{B}\|^{2m} \left(\alpha(C) + \frac{\|\mathcal{B}\|}{K^2} \right),
\end{aligned}$$

where $K = \prod_{i=1}^C |\Phi_{d_i, k_i}(N)|$, and the third line is an application of Theorem 3.17.

Thus,

$$\begin{aligned}
& \left(\mathbf{E}_X \left[|f_3(x_1)| \prod_{j=1}^m 1_{[\mathcal{B}(L_j(X))=b_j]} \right] \right)^2 \leq \varepsilon^2 \left(\frac{2}{\|\mathcal{B}\|} \right) \left(\alpha(C) + \frac{\|\mathcal{B}\|}{K^2} \right) \\
& \leq 2\varepsilon^2 \left(\alpha(C) + \frac{1}{K^2} \right),
\end{aligned}$$

so each term in our original sum that involves f_3 has magnitude at most

$$\sqrt{2\varepsilon^2 \left(\alpha(C) + \frac{1}{K^2} \right)} \leq \varepsilon \sqrt{2 \left(\alpha(C) + \frac{1}{K^2} \right)}.$$

Finally, we bring everything together. Note that $K \leq \prod_{i=1}^C |\Phi_{d_i, k_i}(N)| \leq \|\mathcal{B}\|^m \leq 2^{dCm}$, so $\alpha(C) \leq \frac{1}{K^2}$. Setting $\eta(C) = \left(\frac{\zeta}{3}\right)^m 2^{-dCm-3}$, $\varepsilon_0 = \frac{1}{16} \left(\frac{\zeta}{2}\right)^m$, we have

$$\begin{aligned}
& \mathbf{E}_X \left[\prod_{j=1}^m f(L_j(X)) \right] \geq \left(\frac{1}{K} - \alpha(C) \right) \zeta^m - 2^{m+\frac{1}{2}} \varepsilon \sqrt{\alpha(C) + \frac{1}{K^2}} - 3^m \eta(C) \|\mathcal{B}\| \\
& \geq \frac{1}{2K} \zeta^m - 2^{m+1} \varepsilon \frac{1}{K} - \frac{1}{8K} \zeta^m \\
& \geq \zeta^m \frac{1}{4K} \geq \frac{\zeta^m}{4} \frac{1}{\|\mathcal{B}\|^m}.
\end{aligned}$$

So, there are at least $\frac{\zeta^m}{4} \frac{(2^n)^\ell}{\|\mathcal{B}\|^m}$ linear maps ι such that $\iota(N) \subseteq M$. At most $\ell(2^n)^{\ell-1}2^{\ell-1}$ such linear maps are not injections. When n is sufficiently large compared to ℓ and C , this is negligible compared to the number of linear maps we obtained, so for some choice of ν , the number of injections ι taking N into M is at least $\frac{\zeta^m}{5} \frac{(2^n)^\ell}{\|\mathcal{B}\|^m}$ for large enough n . Since N has at most 2^{ℓ^2} automorphisms, there are at least $\frac{\zeta^m}{5 \cdot 2^{\ell^2}} \frac{(2^n)^\ell}{\|\mathcal{B}\|^m}$ distinct copies of N in M . Taking $\beta = \frac{\zeta^m}{5 \cdot 2^{\ell^2}}$, this concludes the proof of the Counting Lemma. \square

4 Basic Applications

4.1 The Removal Lemma

As a first application of our Counting Lemma, we give a simple proof of a removal lemma for matroids. This removal lemma is a special case of a removal lemma for linear equations that appears in [11] and [13], in both cases proven using a hypergraph removal lemma.

We say that a matroid $M \subseteq \mathbb{F}^{r(M)} \setminus \{0\}$ is ε -far from being N -free if for any matroid $M' \subseteq \mathbb{F}^{r(M)} \setminus \{0\}$ that does not contain a copy of N , $\mathbf{E}_x[|1_M - 1_{M'}|] \geq \varepsilon$.

Theorem 4.1 (Removal Lemma). *For any $\zeta > 0$ and matroid N there is a $\alpha > 0$ such that if a matroid M of sufficiently high rank $r(M)$ is ζ -far from being N -free, then M contains at least $\alpha(2^{r(M)})^{r(N)}$ copies of N .*

Proof. Let $\zeta' = \frac{\zeta}{4}$, and let $d = |N| - 2$. By the Counting Lemma, there exist $\beta, \eta, \varepsilon, r, \nu$ such that if the reduced matroid $R = R_{\varepsilon, \zeta'}$ given by an (η, r, d) -regular partition of a matroid M with corresponding factor \mathcal{B} contains a copy of N and $r(M) \geq \nu(|\mathcal{B}|)$, then M contains at least $\beta \frac{(2^{r(M)})^{r(N)}}{\|\mathcal{B}\|^{|N|}}$ copies of N . Fix such a choice of $\beta, \eta, \varepsilon, r, \nu$.

Let M be a matroid of rank n . Suppose that M is ζ -far from being N -free.

By Theorem 3.11, we have a $(\varepsilon\zeta'^{1/2}, \eta, r, d)$ -regular partition of M , $1_M = f_1 + f_2 + f_3$, whose corresponding factor \mathcal{B} has complexity at most C , where C depends on only $\zeta', \varepsilon, \eta, r, d$, i.e. depends on only ζ, N . Let $M' = M \cap R_{\varepsilon, \zeta'}$.

The only elements in $M \setminus M'$ are either

- (i) In an atom b of \mathcal{B} such that $E[|f_3(x)|^2 \mid x \in b] > \varepsilon^2$, or
- (ii) In an atom b of \mathcal{B} such that $E[f(x) \mid x \in b] < \zeta'$.

Let S be the subset of \mathbb{F}^n contained in atoms b of \mathcal{B} such that $E[|f_3(x)|^2 \mid x \in b] > \varepsilon^2$. Then by condition (iv) of Theorem 3.11,

$$\varepsilon^2 \zeta' \geq \|f_3\|_2^2 = E_x[|f_3(x)|^2] \geq \frac{|S|}{2^n} \varepsilon^2,$$

so $|S| \leq \zeta' 2^n$. Likewise, let T be the subset of \mathbb{F}^n contained in atoms b of \mathcal{B} such that $E[f(x) \mid x \in b] < \zeta'$. Then $|T \cap M| < \zeta' |T| \leq \zeta' 2^n$.

So, $E_x|1_M - 1_{M'}| \leq \frac{|S|+|T|}{2^n} < 2\zeta' = \frac{\zeta}{2}$. Since M is ζ -far from being N -free, M' cannot be N -free, so M' contains a copy of N . If we take $n \geq \nu(C)$, then the Counting Lemma argument above yields that M contains at least $\beta \frac{(2^n)^{r(N)}}{\|\mathcal{B}\|^{|N|}} \geq \frac{1}{2^{dC|N|}} (2^n)^{r(N)}$ copies of N , so we are done by taking $\alpha = \frac{1}{2^{dC|N|}}$. \square

4.2 The Doubling Lemma and the Geometric Erdős-Stone Theorem

We extend the argument in the proof of the removal lemma to prove a simple result we will call the *doubling lemma*. To state it, we define the double of a matroid.

Definition 4.2. Let N be a matroid of rank ℓ . Define its *double* $2N$ to be the matroid of rank $\ell + 1$ consisting of the union of N with $\{x + v | x \in N\}$, where v is a nonzero element not contained in the span of the elements of N . So, for example, $2BB(n, c) = BB(n + 1, c)$. The matroid $2^k N$ is the result of starting with N and doubling k times.

The matroid $2^k N$ corresponds to the result of replacing each element of N with an affine cube of dimension k . Note that if there is a homomorphism from N to R , then for any $k > 1$ there is a homomorphism from $2^k N$ to R because we can simply first contract each of the affine hypercubes into a point.

Lemma 4.3 (Doubling Lemma). *For any $\alpha > 0$ and matroid N there exists $\alpha' > 0$ such that for sufficiently large n , if a matroid $M \subseteq \mathbb{F}^n \setminus 0$ contains at least $\alpha(2^n)^{r(N)}$ copies of N , then it contains at least $\alpha'(2^n)^{r(N)+1}$ copies of $2N$.*

Proof. Let M be a matroid of rank n , and suppose that M contains at least $\alpha(2^n)^{r(N)}$ copies of N . Any element $v \in M$ can be contained in at most $|N|(2^n)^{r(N)-1}$ copies of N , so if M' is a subset of M with $|M'| < \frac{\alpha}{|N|} 2^n$, then $M \setminus M'$ contains a copy of N . So M is $\frac{\alpha}{|N|}$ -far from being N -free. Set $\zeta := \frac{\alpha}{|N|}$.

Let $\zeta' = \frac{\zeta}{4}$, $d = 2|N| - 2$. By the Counting Lemma, there exist $\beta, \eta, \varepsilon, r, \nu$ such that if there exists a homomorphism from $2N$ to the reduced matroid $R = R_{\varepsilon, \zeta'}$ given by an (η, r, d) -regular partition of a matroid M with corresponding factor \mathcal{B} , and $r(M) \geq \nu(|\mathcal{B}|)$, then M contains at least $\beta \frac{(2^{r(M)})^{r(N)+1}}{\|\mathcal{B}\|^{|2N|}}$ copies of $2N$. Fix such a choice of $\beta, \eta, \varepsilon, r, \nu$.

By Theorem 3.11, we have a $(\varepsilon\zeta'^{1/2}, \eta, r, d)$ -regular partition of M , $1_M = f_1 + f_2 + f_3$, whose corresponding factor \mathcal{B} has complexity at most C , where C depends on only $\zeta', \varepsilon, \eta, r, d$, i.e. depends on only ζ, N . Let $M' = M \cap R_{\varepsilon, \zeta'}$.

By the same argument as in the proof of the removal lemma, M' contains a copy of N . Thus there is a homomorphism from $2N$ to $R_{\varepsilon, \zeta'}$.

So, if we take $n \geq \nu(C)$, by the Counting Lemma argument above, M contains at least $\beta \frac{(2^n)^{r(N)+1}}{\|\mathcal{B}\|^{|2N|}} \geq \frac{1}{2^{2dC|N|}} (2^n)^{r(N)+1}$ copies of N , so we are done by taking $\alpha' = \frac{1}{2^{2dC|N|}}$. \square

Remark. We could also have directly gotten a (nondegenerate) copy of $2N$ in $R_{\varepsilon, \zeta'}$ by using an extension of the Chevalley-Waring Theorem to prime power moduli, such as Theorem B in [12].

One particular special case of this result is of interest.

Corollary 4.4. *For any $\alpha > 0$ and positive integers s, t there exists $\alpha' > 0$ such that for sufficiently large n , if a matroid $M \subseteq \mathbb{F}^n \setminus 0$ contains at least $\alpha(2^n)^s$ copies of $PG(s-1, 2)$, then it contains at least $\alpha'(2^n)^{s+t}$ copies of $BB(s+t, s)$.*

Corollary 4.4 should be compared with the following analogous graph-theoretic lemma, used in the proof of results on the chromatic threshold in [2].

Lemma 4.5 ([1, Lemma 7]). *For every r, s and $\varepsilon > 0$ there exists $\delta = \delta_{r,s}(\varepsilon) > 0$ such that the following holds for sufficiently large n . If the n -vertex graph G contains at least εn^r copies of K_r , then G contains at least $\delta_{r,s}(\varepsilon) n^{rs}$ copies of $K_r(s)$.*

Along the same lines, our Counting Lemma gives a short proof of the Geometric Erdős-Stone theorem, Theorem 2.9, using the Bose-Burton theorem, Theorem 2.10, analogous to the proof of the Erdős-Stone theorem using Turán's theorem.

Proof of Theorem 2.9. Let $r(M) = n$, $\chi(N) = c$, and $r(N) = \ell$, and suppose $|M| \geq (1 - 2^{1-c} + \zeta)2^n$. It suffices to show that if n is sufficiently large, M must contain a copy of $BB(\ell, c)$, so without loss of generality $N = BB(\ell, c)$.

Let $\zeta' = \frac{\zeta}{4}$, and let $d = |N| - 2$. By the Counting Lemma, there exist $\beta, \eta, \varepsilon, r, \nu$ such that if there is a homomorphism from N to the reduced matroid $R = R_{\varepsilon, \zeta'}$ given by an (η, r, d) -regular partition of a matroid M with corresponding factor \mathcal{B} , and $r(M) \geq \nu(|\mathcal{B}|)$, then M contains at least $\beta \frac{(2^n)^t}{\|\mathcal{B}\|^{|N|}} > 0$ copies of N . Fix such a choice of $\beta, \eta, \varepsilon, r, \nu$.

By Theorem 3.11, we have a $(\varepsilon \zeta'^{1/2}, \eta, r, d)$ -regular partition of M , $1_M = f_1 + f_2 + f_3$, whose corresponding factor \mathcal{B} has complexity at most C , where C depends on only $\zeta', \varepsilon, \eta, r, d$, i.e. depends on only ζ, N . Let $M' = M \cap R_{\varepsilon, \zeta'}$. As in the proof of the Removal Lemma, we see that $E_x|1_M - 1_{M'}| < \frac{\zeta}{2}$, so that $|M'| \geq (1 - 2^{1-c} + \frac{\zeta}{2})2^n$. By the Bose-Burton theorem, $M' \subseteq R_{\varepsilon, \zeta'}$ contains a copy of $PG(c-1, 2)$. So there is a homomorphism from $N = 2^{\ell-c}PG(c-1, 2)$ to $R_{\varepsilon, \zeta'}$, and thus the Counting Lemma gives us at least one copy of N in M , as desired. \square

5 Applications to the Critical Threshold Problem

The strong decomposition theorem and our Counting Lemma allow us to extend the arguments using Green's regularity lemma in [7] and [6] to address more general cases of Conjecture 2.13. To illustrate the approach we take, we first extend the argument in [7] to the case of $N_{\ell, 2, 1}$. This case is simple enough that it suffices to only use Green's regularity lemma, but we phrase it in terms of the

$d = 1$ case of the strong decomposition theorem to highlight the similarity with our approach to a more general case, in the next section.

5.1 Verifying the Conjecture for $N_{\ell,2,1}$

Proposition 5.1. $\theta(N_{\ell,2,1}) = \frac{1}{4}$.

Proof. Fix $\delta > 0$.

Let $\zeta = \frac{\delta}{2}$, and let $d = 1$. By the Counting Lemma, there exist $\beta, \eta, \varepsilon, r, \nu$ such that if the reduced matroid $R = R_{\varepsilon, \zeta}$ given by an (η, r, d) -regular partition of a matroid M with corresponding factor \mathcal{B} contains a copy of $PG(1, 2)$ and $r(M) \geq \nu(|\mathcal{B}|)$, then M contains at least $\beta \frac{(2^{r(M)})^2}{\|\mathcal{B}\|^3}$ copies of $PG(1, 2)$. Fix such a choice of $\beta, \eta, \varepsilon, r, \nu$.

Let M be a matroid of rank $n \geq \nu(|\mathcal{B}|)$ with $|M| \geq (\frac{1}{4} + \delta)2^n$. By Theorem 3.11, we have an $(\varepsilon\zeta^{1/2}, \eta, r, 1)$ -regular partition of M , $1_M = f_1 + f_2 + f_3$, whose corresponding factor \mathcal{B} has complexity at most C , where C depends only on $\zeta, \varepsilon, \eta, r, d$, i.e. depends only on δ . Let $R = R_{\varepsilon, \zeta}$ and let $D = R_{\varepsilon, \frac{1}{2} + \zeta}$. We have two cases, depending on whether D is nonempty.

Case 1: D is nonempty; that is, for some atom b of \mathcal{B} , $|b \cap M| \geq (\frac{1}{2} + \zeta)|b|$ and $\mathbf{E}[|f_3(x)|^2 \mid x \in b] \leq \varepsilon^2$. Since \mathcal{B} is a factor defined by linear polynomials, if h is any element of the atom b_0 containing 0, then shifting by h preserves each atom. If $\chi(M) > C$, then there exists such an element h in $M \cap b_0$. Consider the submatroid $M_h = \{w \in M \mid w + h \in M\}$. Since $|M \cap b| \geq (\frac{1}{2} + \zeta)|b|$, $|M_h \cap b| \geq \zeta|b| \geq \frac{\zeta}{\|\mathcal{B}\|}2^n$. By the Density Hales-Jewett Theorem, for sufficiently large n , M_h contains a copy A of $AG(m-1, 2)$. Then $A \cup (A+h) \cup \{h\}$ contains a copy of $N_{\ell,2,1}$. So, either $\chi(M) \leq C$ or M contains a copy of $N_{\ell,2,1}$, as desired.

Case 2: D is empty.

Then for each atom b of \mathcal{B} , either $\mathbf{E}[|f_3(x)|^2 \mid x \in b] > \varepsilon^2$ or $|b \cap M| < \frac{1}{2} + \zeta$. Since $|f_3|_{L^2} \leq \varepsilon\zeta^{1/2}$, the former is true for less than a fraction ζ of the atoms b . We can use this to give a lower bound on the size of R . Indeed, we have

$$(\frac{1}{4} + \delta)2^n \leq |M| < (\zeta 2^n + |D|) \cdot 1 + (|R| - |D|) \cdot (\frac{1}{2} + \zeta) + (2^n - |R|) \cdot \zeta,$$

so $|R| > (\frac{1}{2} + 2(\delta - 2\zeta))2^n > \frac{1}{2}2^n$.

By Theorem 2.10, R contains a copy of $PG(1, 2)$. So, by the Counting Lemma, M contains at least $\beta \frac{(2^n)^2}{\|\mathcal{B}\|^3}$ copies of $PG(1, 2)$. Then some element h is part of at least $\beta \frac{2^n}{\|\mathcal{B}\|^3}$ copies of $PG(1, 2)$, so M_h , as defined in Case 1, has density at least $\frac{\beta}{\|\mathcal{B}\|^3}$. Applying the Density Hales-Jewett Theorem again gives a copy of $AG(\ell-1, 2)$ in M_h , and thus a copy of $N_{\ell,2,1}$ in M , as desired. \square

5.2 Conditional Result on $N_{\ell,c,1}$

The case $c > 2$ is more difficult to address because using the Counting Lemma now requires polynomial factors of higher degrees, with which the construction

of M_h from before does not interact in as simple a manner. To make progress on this case, we will need to make additional assumptions on the polynomial factors that we obtain.

Here, we will illustrate our techniques by verifying Conjecture 2.13 in the case where $N = N_{\ell,c,1}$, assuming the following conjecture on the regularity of factors formed by intersecting two shifts of a given polynomial factor.

Conjecture 5.2. *For any nondecreasing function $r' : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ and parameters δ, η, r, d , there exists a constant K such that the polynomial factor \mathcal{B} of $V = \mathbb{F}^n$ given by Theorem 3.11 can be chosen to satisfy the following property. Let \mathcal{B} be defined by polynomials P_1, \dots, P_C , let $\Delta_h P_i(x) = P_i(x+h) - P_i(x) - P_i(h)$ for $1 \leq i \leq C$, and let S be the set of values $h \in V$ for which the factor defined by polynomials $P_1, \dots, P_C, \Delta_h P_1, \dots, \Delta_h P_C$ is r' -regular. Then S contains a subspace of V of codimension at most K .*

We remark that we are not confident in the truth of this conjecture. However, we believe that the arguments that follow are useful in demonstrating some new techniques that seem promising for future work on the critical threshold problem.

Theorem 5.3. *Assuming Conjecture 5.2 holds, $\theta(N_{\ell,c,1}) = 1 - 3 \cdot 2^{-c}$ for all $c \geq 2$.*

Proof. Assume $c \geq 3$, since the case $c = 2$ is Proposition 5.1. Fix $\delta > 0$.

Let $N = N_{\ell,c,1}$, and let $N^* = BB(\ell - 1, c - 1)$, so $N = N^* \cup (N^* + v) \cup \{v\}$ for some element v . Let $\zeta = \frac{\delta}{2}$ and let $d = |N| - 2$. By the Counting Lemma, there exist $\beta_1, \eta_1, \varepsilon_1, r_1, \nu_1$ such that if the reduced matroid $R = R_{\varepsilon_1, \zeta}$ given by an (η_1, r_1, d) -regular partition of a matroid M with corresponding factor \mathcal{B} contains a copy of N and $r(M) \geq \nu_1(|\mathcal{B}|)$, then M contains at least $\beta_1 \frac{(2^{r(M)})^{r(N)}}{\|\mathcal{B}\|^{|\mathcal{B}|}}$ copies of N . Let $\beta_2, \eta_2, \varepsilon_2, r_2, \nu_2$ be parameters we will pick later, and define $\beta = \min(\beta_1, \beta_2)$, $\eta(C) = \min(\eta_1(C), \eta_2(C))$, $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$, $r(C) = \max(r_1(C), r_2(C))$, $\nu(C) = \max(\nu_1(C), \nu_2(C))$. Let $\varepsilon' = \frac{1}{2}\varepsilon\zeta^{1/2}$.

Let M be a matroid of rank $n \geq \nu(|\mathcal{B}|)$ with $|M| \geq (1 - 3 \cdot 2^{-c} + \delta)2^n$. By Theorem 3.11, we have a $(\frac{1}{2}\varepsilon'\zeta^{1/2}, \eta, r, 1)$ -regular partition of M , $1_M = f_1 + f_2 + f_3$, whose corresponding factor \mathcal{B} has complexity at most C , where C depends only on $\zeta, \varepsilon, \eta, r, d$, i.e. depends only on $\delta, |N|$. Assuming Conjecture 5.2, for some K, r_2 depending on only N and d , we can pick \mathcal{B} such that the set S of elements h for which \mathcal{B}_h is r'_2 -regular contains a subspace W of $V = \mathbb{F}^n$ of codimension at most K as long as \mathcal{B} is r_2 -regular, where r'_2 is another parameter to be determined later. Let $R = R_{\varepsilon', \zeta}$ and let $D = R_{\varepsilon', \frac{1}{2} + \zeta}$. We have two cases, depending on the density of D .

Case 1: $|D| \leq (1 - 2^{2^{-c}} + \zeta)2^n$.

For each atom b of \mathcal{B} , either $\mathbf{E}[|f_3(x)|^2 \mid x \in b] > \varepsilon'^2$, $b \subseteq D$, or $|b \cap M| < \frac{1}{2} + \zeta$. Since $\|f_3\|_{L^2} \leq \frac{1}{2}\varepsilon'\zeta^{1/2}$, the atoms b for which the first is true span in total less than a fraction $\frac{\zeta}{4}$ of the elements of \mathbb{F}^n . We can use this to give a lower bound on the size of R . Indeed, we have

$$(1 - 3 \cdot 2^{-c} + \delta)2^n \leq |M| < \left(\frac{\zeta}{4}2^n + |D|\right) \cdot 1 + (|R| - |D|) \cdot \left(\frac{1}{2} + \zeta\right) + (2^n - |R|) \cdot \zeta,$$

so

$$\begin{aligned}
|R| &> 2 \left((1 - 3 \cdot 2^{-c} + \delta)2^n - |D| \left(\frac{1}{2} - \zeta \right) - \frac{5}{4}\zeta 2^n \right) \\
&\geq \left(1 - 2^{1-c} + 2\left(\delta - \frac{5}{4}\zeta\right) - \zeta + 2\zeta(1 - 2^{2-c} + \zeta) \right) 2^n \\
&\geq (1 - 2^{1-c} + \zeta)2^n.
\end{aligned}$$

By the Bose-Burton Theorem, R contains a copy of $PG(c-1, 2)$, and thus there exists a homomorphism from $2^{\ell-c}PG(c-1, 2)$ to R . Since N is a submatroid of $2^{\ell-c}PG(c-1, 2)$, there is a homomorphism from N to R . So, since $R \subseteq R_{\varepsilon, \zeta}$, by the Counting Lemma, M contains a positive number of copies of N , as desired.

Case 2: $|D| > (1 - 2^{2-c} + \zeta)2^n$.

We will introduce a few new technical tools to deal with this case.

Assume without loss of generality that $|\mathcal{B}| = C$. Given an element $h \in V$, let \mathcal{B}_h be the factor defined by the polynomials $P_1, \dots, P_C, \Delta_h P_1, \dots, \Delta_h P_C$. Note that the indicator function for $D \cap (D + h)$ is constant on each atom of \mathcal{B}_h .

We represent N^* as a system of linear forms on $\ell-1$ variables, $\{L_1, \dots, L_m\}$, where without loss of generality $\{L_1, \dots, L_{2^{c-1}-1}\}$ forms a copy of $PG(c-1, 2)$. Let $M_h = \{w \in M \mid w + h \in M\}$. For an appropriately chosen h , we will derive a lower bound for the number of copies of N^* contained in M_h . If $h \in M$, we will then end up with a copy of N in M .

Let $g(x) = f(x)f(x+h)$ be the indicator function for M_h , so the expression we wish to give a lower bound for is

$$\mathbf{E}_{X \in (\mathbb{F}^n)^{\ell-1}} \left[\prod_{j=1}^m f(L_j(X))f(L_j(X) + h) \right].$$

Note that

$$\begin{aligned}
f(x)f(x+h) &= (1 - f(x))(1 - f(x+h)) + (f(x) + f(x+h) - 1) \\
&= ((1 - f(x))(1 - f(x+h)) + f_1(x) + f_1(x+h) - 1) \\
&\quad + (f_2(x) + f_2(x+h)) + (f_3(x) + f_3(x+h)).
\end{aligned}$$

Let $g_1(x) = (1 - f(x))(1 - f(x+h)) + f_1(x) + f_1(x+h) - 1$, $g_2(x) = f_2(x) + f_2(x+h)$, and $g_3(x) = f_3(x) + f_3(x+h)$. So, $g(x) = g_1(x) + g_2(x) + g_3(x)$.

Similarly to the approach in the proof of the Counting Lemma, we will restrict to counting within certain “good” atoms of \mathcal{B}_h . Specifically, suppose we have a point $Y \in (\mathbb{F}^n)^{\ell-1}$ such that $L_1(Y), \dots, L_m(Y)$ are in atoms $\tilde{b}_1, \dots, \tilde{b}_m$ of \mathcal{B}_h contained in $D \cap (D+h)$ such that $\mathbf{E}[|f_3(x)|^2 \mid x \in \tilde{b}_j], \mathbf{E}[|f_3(x+h)|^2 \mid x \in \tilde{b}_j]$ are at most ε^2 for $1 \leq j \leq m$. Then

$$\begin{aligned}
& \mathbf{E}_X \left[\prod_{j=1}^m f(L_j(X)) f(L_j(X) + h) \right] \geq \mathbf{E}_X \left[\prod_{j=1}^m f(L_j(X)) f(L_j(X) + h) 1_{[\mathcal{B}_h(L_j(X))=\tilde{b}_j]} \right] \\
& = \sum_{(i_1, \dots, i_m) \in [1,3]^m} \mathbf{E}_X \left[\prod_{j=1}^m g_{i_j}(L_j(X)) 1_{[\mathcal{B}_h(L_j(X))=\tilde{b}_j]} \right].
\end{aligned}$$

To handle the terms where $i_j = 2$ for some j , we use the following lemma.

Lemma 5.4. *Let $s \geq 1$. For any (nonclassical) polynomial P of degree at most s , constant $\beta \in \mathbb{T}$, and function $g : \mathbb{F}^n \rightarrow \mathbb{C}$, we have*

$$\|g 1_{P(x)=\beta}\|_{U^{s+1}} \leq \|g\|_{U^{s+1}}.$$

Proof. Let P have depth k . We have

$$g(x) 1_{P(x)=\beta}(x) = \frac{1}{2^{k+1}} \sum_{\lambda=0}^{2^{k+1}-1} g(x) \mathbf{e}(\lambda(P(x) - \beta)).$$

Since $s+1 \geq 2$, the triangle inequality holds for the Gowers norm, so

$$\begin{aligned}
\|g 1_{P(x)=\beta}\|_{U^{s+1}} &= \left\| \frac{1}{2^{k+1}} \sum_{\lambda=0}^{2^{k+1}-1} g(x) \mathbf{e}(\lambda(P(x) - \beta)) \right\|_{U^{s+1}} \\
&\leq \frac{1}{2^{k+1}} \sum_{\lambda=0}^{2^{k+1}-1} \|g(x) \mathbf{e}(\lambda(P(x) - \beta))\|_{U^{s+1}} = \|g\|_{U^{s+1}},
\end{aligned}$$

since $\|g \cdot P(x)\|_{U^{s+1}} = \|g\|_{U^{s+1}}$ when P has degree $\leq s$. \square

Repeated application of Lemma 5.4, followed by the triangle inequality, gives that

$$\|g_2(x) 1_{[\mathcal{B}_h(x)=\tilde{b}_j]}\|_{U^{d+1}} \leq \|g_2\|_{U^{d+1}} \leq 2\|f_2\|_{U^{d+1}},$$

for $1 \leq j \leq m$. So, since $\max_x |g_i(x)| \leq 2$ for $1 \leq i \leq 3$, applying Lemma 3.15 on $\{\frac{1}{2}g_{i_j}\}_{j=1}^m$ gives

$$\begin{aligned}
& \left| \mathbf{E}_X \left[\prod_{j=1}^m g_{i_j}(L_j(X)) 1_{[\mathcal{B}_h(L_j(X))=\tilde{b}_j]} \right] \right| \leq 2^m \min_{1 \leq j \leq m} \left\| \frac{1}{2} g_{i_j}(x) 1_{[\mathcal{B}_h(x)=\tilde{b}_j]} \right\|_{U^{d+1}} \\
& \leq 2^m \left\| \frac{1}{2} g_2(x) 1_{[\mathcal{B}_h(x)=\tilde{b}_j]} \right\|_{U^{d+1}} \leq 2^m \|f_2\|_{U^{d+1}} \leq 2^m \eta(|\mathcal{B}|),
\end{aligned}$$

for each term where at least one of the i_j is 2.

Our probability is thus at least

$$\sum_{(i_1, \dots, i_m) \in \{1, 3\}^m} \mathbf{E}_X \left[\prod_{j=1}^m g_{i_j}(L_j(X)) 1_{[\mathcal{B}_h(L_j(X)) = \tilde{b}_j]} \right] - 6^m \eta(|\mathcal{B}|).$$

When $x \in \tilde{b}_j$, we have $f_1(x), f_1(x+h) \geq \frac{1}{2} + \zeta$, so $g_1(x) = (1 - f(x))(1 - f(x+h)) + f_1(x) + f_1(x+h) \geq 2\zeta$. The argument after this is exactly as in the proof of the Counting Lemma, assuming that \mathcal{B}_h is r' -regular for r' sufficiently large. We get a lower bound on the main term by applying Theorem 3.18, and use Cauchy-Schwarz followed by Lemma 3.20 and Theorem 3.17 to deal with the terms involving g_3 but not g_2 . By that argument, we are guaranteed at least $\beta \frac{(2^n)^{\ell-1}}{\|\mathcal{B}\|^{N^*}}$ copies of N^* in M_h as long as $\beta \leq \beta_2$, $\eta(|\mathcal{B}_h|) \leq \eta_2(|\mathcal{B}_h|)$, $\varepsilon \leq \varepsilon_2$, $r'(|\mathcal{B}_h|) \geq r'_2(|\mathcal{B}_h|)$, and $n \geq \nu_2(|\mathcal{B}_h|)$ for some parameters $\beta_2, \eta_2, \varepsilon_2, r'_2, \nu_2$ depending only on N, d . At most $O_\ell((2^n)^{\ell-2})$ of these copies intersect themselves when shifted by h , so without loss of generality ν_2 is big enough such that a copy A exists where A is disjoint from $A+h$, and so $A \cup (A+h) \cup \{h\}$ forms a copy of N in M .

It remains to show that, when $\chi(M)$ is sufficiently large, there exists an $h \in M$ where \mathcal{B}_h is r'_2 -regular and a point $Y \in (\mathbb{F}^n)^{\ell-1}$ such that $L_1(Y), \dots, L_m(Y)$ are in atoms $\tilde{b}_1, \dots, \tilde{b}_m$ of \mathcal{B}_h contained in $D \cap (D+h)$ such that $\mathbf{E}[|f_3(x)|^2 \mid x \in \tilde{b}_j], \mathbf{E}[|f_3(x+h)|^2 \mid x \in \tilde{b}_j]$ are at most ε^2 for $1 \leq j \leq m$. By Proposition 3.10 we can, without loss of generality, require r_2, r'_2 to be large enough (depending only on ζ, d, C) so that \mathcal{B} and \mathcal{B}_h are $\frac{\zeta}{16\|\mathcal{B}\|}$ -uniform. In this case, by Corollary 3.19, every potential atom of \mathcal{B}_h consistent with the defining polynomials is nonempty, and in fact has size at least $\left(\frac{1}{\|\mathcal{B}_h\|} - \frac{\zeta}{16\|\mathcal{B}_h\|}\right) 2^n$.

Recall that W is a subspace of $V = \mathbb{F}^n$ of codimension at most K such that \mathcal{B}_h is r'_2 -regular for all $h \in W$. For any $h \in W$ and $i \in [1, C]$, by our choice of r'_2 , $\Delta_h P_i$ attains all values in $\frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z}$, where k'_i is the depth of $\Delta_h P_i$. So, $P_i(x+h) - P_i(x) = \Delta_h P_i(x) + P_i(h)$ attains all values in a coset of $\frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z}$ in $\frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z} \subset \mathbb{T}$. We will use the following proposition to narrow down our choice of h to a slightly smaller subspace, in order to ensure that $P_i(h) \in \frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z}$.

Proposition 5.5. *Let P be a (nonclassical) polynomial of depth at most k on $V = \mathbb{F}^n$ and let W be a subspace of V such that for all $h \in W$, $\Delta_h P$ attains all values in $\frac{1}{2^{k'_h+1}}\mathbb{Z}/\mathbb{Z}$, where k'_h is the depth of $\Delta_h P$. Then there exists a subspace W' of W with codimension at most $k+1$ in W , such that for $h \in W'$, $P(h) \in \frac{1}{2^{k'_h+1}}\mathbb{Z}/\mathbb{Z}$.*

Proof. Without loss of generality let P have depth k , so for each h , $\Delta_h P$ has depth $k_h \leq k$. We use induction on i to show that, for $0 \leq i \leq k+1$, the set of h for which $P(x+h) - P(x)$ attains no value in $\frac{1}{2^{k+1-i}}\mathbb{Z}/\mathbb{Z}$ is contained in the complement of a subspace of W of codimension i . This gives a subspace of codimension at most $k+1$ where $P(x+h) - P(x)$ attains the value 0, which implies $P(h) \in \frac{1}{2^{k'_h+1}}\mathbb{Z}/\mathbb{Z}$.

The statement is trivial for $i = 0$ since by assumption, $\Delta_h P$ attains some value in $\frac{1}{2^{k'_h+1}}\mathbb{Z}/\mathbb{Z} \subseteq \frac{1}{2^{k+1}}\mathbb{Z}/\mathbb{Z}$ for all $h \in W$.

Assume the statement is true for $i = j \leq k$, so there is some subspace $W_j \leq W$ of codimension j such that for all $h \in W_j$, $P(x+h) - P(x)$ attains some value in $\frac{1}{2^{k+1-j}}\mathbb{Z}/\mathbb{Z}$. Let Z be the subset of W_j consisting of the values of h for which $P(x+h) - P(x)$ attains no value in $\frac{1}{2^{k+1-(j+1)}}\mathbb{Z}/\mathbb{Z}$. Since the values attained by $P(x+h) - P(x)$ form a coset of some subgroup $\frac{1}{2^{k'_h+1}}\mathbb{Z}/\mathbb{Z}$ of \mathbb{T} , and this coset must contain some element of $\frac{1}{2^{k+1-j}}\mathbb{Z}/\mathbb{Z}$ but avoid 0, we must have $k'_h < k - j$. So, for all $h \in Z$, all values attained by $P(x+h) - P(x)$ are contained in $\frac{1}{2^{k+1-j}} + \frac{1}{2^{k-j}}\mathbb{Z}/\mathbb{Z}$.

Assume for the sake of contradiction that Z contains an odd circuit, so for some $h_1, \dots, h_t \in Z$, $\sum_{s=1}^t h_s = 0$, where t is odd. But

$$0 = P\left(x + \sum_{s=1}^t h_s\right) - P(x) = \sum_{s=1}^t \left(P\left(x + \sum_{u=1}^s h_u\right) - P\left(x + \sum_{u=1}^{s-1} h_u\right) \right),$$

which must take only values in $\sum_{s=1}^t \left(\frac{1}{2^{k+1-j}} + \frac{1}{2^{k-j}}\mathbb{Z}/\mathbb{Z} \right) = \frac{1}{2^{k+1-j}} + \frac{1}{2^{k-j}}\mathbb{Z}/\mathbb{Z}$. This is a contradiction. So Z contains no odd circuits, and is thus contained in the complement of a hyperplane W_{j+1} of W_j . Thus the statement is true for $i = j + 1$ as well, and by induction is true for all $j \leq k + 1$. This proves the proposition. \square

So, for each P_i we have a subspace W'_i of codimension at most $k_i + 1$ in W on which $P_i(h) \in \frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z}$. Intersecting all of these gives a subspace W' of W of codimension at most $K + \sum_{i=1}^C (k_i + 1) \leq K + dC$ on which $P_i(h) \in \frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z}$ for all i .

If $\chi(M) > K + dC$, we can pick a fixed $h \in M \cap W'$, and let k'_i be the depth of $\Delta_h P_i$ for each i . Since $h \in W'$, and \mathcal{B}_h is sufficiently regular, the range of $P_i(x+h) - P_i(x)$ is $\frac{1}{2^{k'_i+1}}\mathbb{Z}/\mathbb{Z}$, with each value attained approximately equally often.

Let D' be the union of the atoms b of \mathcal{B} that are contained in D and satisfy $\mathbf{E}[|f_3(x+h)|^2 \mid x \in b] \leq \varepsilon'^2$. Since $\|f_3\|_{L^2} \leq \frac{1}{2}\varepsilon'\zeta^{1/2}$, at most a fraction $\frac{\zeta}{4}$ of elements of \mathbb{F}^n are contained in an atom b of \mathcal{B} that does not satisfy $\mathbf{E}[|f_3(x+h)|^2 \mid x \in b] \leq \varepsilon'^2$. Thus $|D'| \geq |D| - \frac{\zeta}{4}2^n > (1 - 2^{2-c} + \frac{3\zeta}{4})2^n$.

We will now prove and use the following proposition, first stated in the introduction, which has a technical statement but a simple proof.

Proposition 1.3. *Let n, c be positive integers, let k_1, \dots, k_n be nonnegative integers, and let $G = \bigoplus_{i=1}^n \frac{1}{2^{k_i+1}}\mathbb{Z}/\mathbb{Z}$. Let H be a subgroup of G . Let M_1, \dots, M_{2^c-1} be subsets of G . Then there exist $H_1, \dots, H_c \in G/H$, cosets of H , such that for $1 \leq i \leq c$,*

$$\frac{1}{|H|} \sum_{x \in [0,1]^{i-1}} \left| M_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}} \cap \left(H_i + \sum_{j=1}^{i-1} x_j H_j \right) \right| \geq \sum_{j=2^{i-1}}^{2^i-1} \frac{|M_j|}{|G|}. \quad (*)$$

Proof. We will choose H_1, \dots, H_c in order, greedily. For $1 \leq i_0 \leq c$, suppose H_1, \dots, H_{i_0-1} have already been chosen such that $(*)$ holds for $1 \leq i \leq i_0 - 1$. Consider a uniformly random choice of $H_{i_0} \in G/H$. Taking an expectation gives

$$\begin{aligned} & \mathbf{E}_{H_{i_0} \in G/H} \frac{1}{|H|} \sum_{x=(x_1, \dots, x_{i_0-1}) \in [0,1]^{i_0-1}} \left| M_{2^{i_0-1} + \sum_{j=1}^{i_0-1} x_j 2^{j-1}} \cap \left(H_{i_0} + \sum_{j=1}^{i_0-1} x_j H_j \right) \right| \\ &= \frac{1}{|H|} \sum_{x=(x_1, \dots, x_{i_0-1}) \in [0,1]^{i_0-1}} \mathbf{E}_{H' \in G/H} \left| M_{2^{i_0-1} + \sum_{j=1}^{i_0-1} x_j 2^{j-1}} \cap H' \right| \\ &= \sum_{j=2^{i-1}}^{2^i-1} \frac{|M_j|}{|G|}, \end{aligned}$$

so for some choice of H_{i_0} , the inequality $(*)$ holds. Continuing in this fashion, we can successfully pick H_i for all $i \in [1, c]$, as desired. \square

Let $G = \bigoplus_{i=1}^C \frac{1}{2^{k_i+1}} \mathbb{Z}/\mathbb{Z}$, and let H be the subgroup of G isomorphic to $\bigoplus_{i=1}^C \frac{1}{2^{k'_i+1}} \mathbb{Z}/\mathbb{Z}$, where each term is a subgroup of the corresponding term in G . To see how this relates to the problem at hand, consider a fixed tuple $b = (b_1, \dots, b_m) \in G^m$, where $b_j = (b_{j,1}, \dots, b_{j,C})$. When \mathcal{B}_h is sufficiently regular, the set of tuples $b' = (b'_1, \dots, b'_m) \in G^m$ for which there is a point $X \in (\mathbb{F}^n)^{\ell-1}$ satisfying $P_i(L_j(X)) = b_{j,i}, P_i(L_j(X) + h) = b'_{j,i}$ for all i, j is exactly those for which b_j, b'_j are in the same coset of H for all j .

Let S be the subset of G consisting of the points $g = (g_1, \dots, g_C)$ such that the atom where $P_i = g_i$ for all i is contained in D' .

By Corollary 3.19 and our choice of r_2 , every atom of \mathcal{B} has size at most $(\frac{1}{\|\mathcal{B}\|} + \frac{\zeta}{16\|\mathcal{B}\|})2^n$. So, since $|D'| > (1 - 2^{2-c} + \frac{3\zeta}{4})2^n$ and $|G| = \|\mathcal{B}\|$, we have $\frac{|S|}{|G|} > \frac{1-2^{2-c}+\frac{3\zeta}{4}}{1+\frac{\zeta}{16}} > 1 - 2^{2-c} + \frac{5}{8}\zeta$.

By Proposition 1.3 with $M_1 = \dots = M_{2^{c-1}-1} = S$, we have cosets H_1, \dots, H_{c-1} of H for which $(*)$ holds for $1 \leq i \leq c-1$. So, for $1 \leq i \leq c-1$,

$$\frac{1}{|H|} \sum_{x \in [0,1]^{i-1}} \left| S \cap \left(H_i + \sum_{j=1}^{i-1} x_j H_j \right) \right| \geq 2^{i-1} \frac{|S|}{|G|} > 2^{i-1} (1 - 2^{2-c} + \frac{5}{8}\zeta).$$

Pick representatives h_1, \dots, h_{c-1} for the cosets H_1, \dots, H_{c-1} . For nonzero $x \in [0,1]^{c-1}$, let $M_{\sum_{j=1}^{c-1} x_j 2^{j-1}} = (S \cap (\sum_{j=1}^{c-1} x_j H_j)) - \sum_{j=1}^{c-1} x_j h_j$. Applying Proposition 1.3 again with H as the group, the trivial group as the subgroup, and the $\{M_j\}_{j=1}^{2^{c-1}-1}$ just defined, we get points $e_1, \dots, e_{c-1} \in H$ such that for

$$1 \leq i \leq c-1,$$

$$\sum_{x \in [0,1]^{i-1}} \left| M_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}} \cap \left\{ e_i + \sum_{j=1}^{i-1} x_j e_j \right\} \right| \geq \sum_{j=2^{i-1}}^{2^i-1} \frac{|M_j|}{|H|}.$$

The right hand side is

$$\begin{aligned} & \sum_{x \in [0,1]^{i-1}} \frac{|M_{2^i + \sum_{j=1}^{i-1} x_j 2^{j-1}}|}{|H|} \\ &= \sum_{x \in [0,1]^{i-1}} \frac{|S \cap (H_i + \sum_{j=1}^{i-1} x_j H_j)|}{|H|} \\ &> 2^{i-1} \left(1 - 2^{2-c} + \frac{5}{8} \zeta \right), \end{aligned}$$

where the last inequality is from the first application of Proposition 1.3. Since $i \leq c-1$, this is strictly greater than $2^{i-1} - 1$. On the other hand, the left hand side is

$$\begin{aligned} & \sum_{x \in [0,1]^{i-1}} \left| M_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}} \cap \left\{ e_i + \sum_{j=1}^{i-1} x_j e_j \right\} \right| \\ &= \sum_{x \in [0,1]^{i-1}} \left| \left(\left(S \cap \left(H_i + \sum_{j=1}^{i-1} x_j H_j \right) \right) - \left(h_i + \sum_{j=1}^{i-1} x_j h_j \right) \right) \cap \left\{ e_i + \sum_{j=1}^{i-1} x_j e_j \right\} \right| \\ &= \left| \left\{ x \in [0,1]^{i-1} \mid (h_i + e_i) + \sum_{j=1}^{i-1} x_j (h_j + e_j) \in \left(S \cap \left(H_i + \sum_{j=1}^{i-1} x_j H_j \right) \right) \right\} \right|, \end{aligned}$$

which is an integer in the interval $[0, 2^{i-1}]$. So this integer must be 2^{i-1} , meaning that for each nonzero $x \in [0,1]^{c-1}$, we have $\sum_{j=1}^{c-1} x_j (h_j + e_j) \in M_{\sum_{j=1}^{c-1} x_j 2^{j-1}} \subseteq S$.

Let $b_{\sum_{j=1}^{c-1} x_j 2^{j-1}} = \sum_{j=1}^{c-1} x_j (h_j + e_j)$ for each nonzero $x \in [0,1]^{c-1}$. We claim that, for $1 \leq i \leq C$, the atoms of G corresponding to $b_1, \dots, b_{2^{c-1}-1}$ are (d_i, k_i) -consistent with the system of linear forms $\{L_1, \dots, L_{2^{c-1}-1}\}$ corresponding to $PG(c-1, 2)$. Indeed, letting $v_j = h_j + e_j = (v_{j,1}, \dots, v_{j,C})$, the function $Q_i(X) = \sum_{j=1}^{c-1} v_{j,i} |x_j|$ is a polynomial of depth at most k_i and degree at most $k_i + 1 \leq d_i$. Thus, there is a copy of $PG(c-1, 2)$ in D' whose elements lie in the atoms $b_1, \dots, b_{2^{c-1}-1}$ of \mathcal{B} .

The next step is to find appropriate atoms $\tilde{b}_1, \dots, \tilde{b}_{2^{c-1}-1}$ of \mathcal{B}_h contained in these atoms of \mathcal{B} . If we fix points $g_1, \dots, g_{2^{c-1}-1} \in H$, where $g_j = (g_{j,1}, \dots, g_{j,C})$, then for $j \in [1, 2^{c-1}-1]$ we can let \tilde{b}_j be the atom of \mathcal{B}_h on which $P_i = b_{j,i}$, $\Delta_h P_i = g_{j,i} + b_{j,i}$ for all $i \in [1, C]$. For $j \in [1, 2^{c-1}-1]$, let S_j be the subset of

H consisting of the points g_j for which $b_j + g_j \in S$ and the atom \tilde{b}_j so defined satisfies $\mathbf{E}[|f_3(x)|^2 \mid x \in \tilde{b}_j], \mathbf{E}[|f_3(x+h)|^2 \mid x \in \tilde{b}_j] \leq \varepsilon^2$.

Since for $b_j \in S$ we have $\mathbf{E}[|f_3(x)|^2 \mid x \in b_j], \mathbf{E}[|f_3(x+h)|^2 \mid x \in b_j] \leq \varepsilon'^2 = \frac{1}{4}\varepsilon^2\zeta$, at most a fraction $\frac{\zeta}{4}$ of the elements of b_j are in an atom \tilde{b}_j of \mathcal{B}_h satisfying $\mathbf{E}[|f_3(x)|^2 \mid x \in \tilde{b}_j] > \varepsilon^2$, and similarly at most a fraction $\frac{\zeta}{4}$ of the elements of b_j are in an atom \tilde{b}_j of \mathcal{B}_h satisfying $\mathbf{E}[|f_3(x+h)|^2 \mid x \in \tilde{b}_j] > \varepsilon^2$. By our choice of r_2 and r'_2 , every atom of \mathcal{B} has size at most $(\frac{1}{\|\mathcal{B}\|} + \frac{\zeta}{16\|\mathcal{B}\|})2^n$ and every atom of \mathcal{B}_h has size at least $(\frac{1}{\|\mathcal{B}_h\|} - \frac{\zeta}{16\|\mathcal{B}_h\|})2^n$, so the number of atoms of \mathcal{B}_h for which one of the two bounds is exceeded is at most $\frac{\zeta \cdot (\frac{1}{\|\mathcal{B}\|} + \frac{\zeta}{16\|\mathcal{B}\|})2^n}{2 \cdot (\frac{1}{\|\mathcal{B}_h\|} - \frac{\zeta}{16\|\mathcal{B}_h\|})2^n} = \frac{\zeta}{2} \frac{1 + \frac{\zeta}{16}}{1 - \frac{\zeta}{16}} |H| \leq \frac{17}{30}\zeta |H|$. That is, $|S_j| \geq |(S - b_j) \cap H| - \frac{17}{30}\zeta |H|$. Applying Proposition 1.3 with H for the group, the trivial group for the subgroup, and sets $S_1, \dots, S_{2^{c-1}-1}$ gives points $p_1, \dots, p_{c-1} \in H$ such that such for $1 \leq i \leq c-1$,

$$\sum_{x \in [0,1]^{i-1}} \left| S_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}} \cap \left\{ p_i + \sum_{j=1}^{i-1} x_j p_j \right\} \right| \geq \sum_{j=2^{i-1}}^{2^i-1} \frac{|S_j|}{|H|}.$$

The right hand side is

$$\begin{aligned} & \sum_{x \in [0,1]^{i-1}} \frac{|S_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}}|}{|H|} \\ & \geq \sum_{x \in [0,1]^{i-1}} \frac{|(S - b_{2^{i-1} + \sum_{j=1}^{i-1} x_j 2^{j-1}}) \cap H| - \frac{17}{30}\zeta |H|}{|H|} \\ & = \sum_{x \in [0,1]^{i-1}} \frac{|S \cap (H_i + \sum_{j=1}^{i-1} x_j H_j)| - \frac{17}{30}\zeta |H|}{|H|} \\ & > 2^{i-1} \left(1 - 2^{2-c} + \frac{5}{8}\zeta - \frac{17}{30}\zeta |H| \right) > 2^{i-1} (1 - 2^{2-c}). \end{aligned}$$

This is greater than $2^{i-1} - 1$. So, by the same argument as before, if for all nonzero $x \in [0,1]^{c-1}$ we let $g_{\sum_{j=1}^{c-1} x_j 2^{j-1}} = \sum_{j=1}^{c-1} x_j p_j$, then $g_j \in S_j$ for all j . Also as before, for $i \in [1, C]$, $(g_{1,i}, \dots, g_{2^{c-1}-1,i})$ is (d'_i, k'_i) -consistent with the system of linear forms $\{L_1, \dots, L_{2^{c-1}-1}\}$. So as long as r'_2 is sufficiently large, by Theorem 3.18, there is a copy of $PG(c-1, 2)$ whose elements are contained in the atoms $\tilde{b}_1, \dots, \tilde{b}_{2^{c-1}-1}$ of \mathcal{B}_h , respectively. Thus there is a homomorphism from N^* into the union of these atoms, so by the Counting Lemma, if $n \geq v_2$ is sufficiently large, there is a copy of N^* in the union of these atoms. By construction, these atoms satisfy the conditions on f_1 and f_3 that we were looking for, so applying the first part of our argument finishes the proof of our conditional result. \square

6 Future Steps

The Counting Lemma has potential for giving simple proofs to other extremal results on matroids whose graph theory analogues are proven using the Szemerédi regularity lemma and its associated counting lemma, including giving short new proofs for known results (as for Theorem 2.9). A search through more such results in extremal graph theory which yield matroid analogues may well prove fruitful.

In terms of the critical threshold problem, if Conjecture 5.2 is shown to be true, then Theorem 5.3 constitutes a proof of a somewhat more general subcase of the $i = 3$ case of Conjecture 2.13 than that given by Theorem 2.11. Regardless of whether Conjecture 5.2 is true (or simple to prove), the methods used in the proof of Theorem 5.3 seem to offer a new approach to the critical threshold problem, that of using a strong form of regularity and attempting to analyze constructions like $M \cap (M + h)$ through counting lemma-type arguments. The same types of arguments seem equally suited for some special subcases of the $i = 4$ case of Conjecture 2.13, and could perhaps be adapted for more general cases as well. Finally, Proposition 1.3, as a simple generalization of the Bose-Burton theorem that makes its proof completely transparent, may be applicable to computing critical thresholds independently of the more sophisticated machinery we have developed.

Acknowledgements

This research was conducted at the University of Minnesota Duluth REU and was supported by NSF grant 1358659 and NSA grant H98230-16-1-0026. The author thanks Joe Gallian for suggesting the problem and for helpful comments on the manuscript. .

References

- [1] ALLEN, P. Dense H -free graphs are almost $(\chi(H) - 1)$ -partite. *Electron. J. Combin.* 17, 1 (2010), R21.
- [2] ALLEN, P., BÖTTCHER, J., GRIFFITHS, S., KOHAYAKAWA, Y., AND MORRIS, R. The chromatic thresholds of graphs. *Adv. Math.* 235 (2013), 261–295.
- [3] BHATTACHARYYA, A., FISCHER, E., HATAMI, H., HATAMI, P., AND LOVETT, S. Every locally characterized affine-invariant property is testable. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing* (2013), ACM, pp. 429–436.
- [4] GEELLEN, J., AND NELSON, P. An analogue of the Erdős-Stone theorem for finite geometries. *Combinatorica* 35, 2 (2015), 209–214.

- [5] GEELEN, J., AND NELSON, P. A density Hales–Jewett theorem for matroids. *J. Combin. Theory Ser. B* 112 (2015), 70–77.
- [6] GEELEN, J., AND NELSON, P. Odd circuits in dense binary matroids. *Combinatorica* (2015), 1–7.
- [7] GEELEN, J., AND NELSON, P. The critical number of dense triangle-free binary matroids. *J. Combin. Theory Ser. B* 116 (2016), 238–249.
- [8] GODDARD, W., AND LYLE, J. Dense graphs with small clique number. *J. Graph Theory* 66, 4 (2011), 319–331.
- [9] GREEN, B. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geom. Funct. Anal.* 15, 2 (2005), 340–376.
- [10] HATAMI, H., HATAMI, P., AND LOVETT, S. General systems of linear forms: equidistribution and true complexity. *Adv. Math.* 292 (2016), 446–477.
- [11] KRÁL’, D., SERRA, O., AND VENA, L. A removal lemma for systems of linear equations over finite fields. *Israel J. Math.* 187, 1 (2012), 193–207.
- [12] SCHANUEL, S. H. An extension of Chevalley’s theorem to congruences modulo prime powers. *J. Number Theory* 6, 4 (1974), 284–290.
- [13] SHAPIRA, A. A proof of Green’s conjecture regarding the removal properties of sets of linear equations. *J. Lond. Math. Soc.* (2010), jdp076.
- [14] TAO, T., AND ZIEGLER, T. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Ann. Comb.* 16, 1 (2012), 121–188.
- [15] TIDOR, J. Dense binary $PG(t-1, 2)$ -free matroids have critical number $t-1$ or t . *arXiv:1508.07278* (2015).